GaN Reliability and Lifetime Projections: Phase 17

Siddhesh Gajare, Ph.D., Duanhui Li, Ph.D., Ricardo Garcia, Angel Espinoza, Jordan Green, Peter Tieu, Ph.D., Christopher Wong, Anthony Nguyen, David Wu, Shengke Zhang, Ph.D.

The rapid adoption of Gallium Nitride (GaN) devices in many diverse applications calls for continued accumulation of reliability statistics and research into the fundamental physics of failure in GaN devices, including integrated circuits (ICs). This Phase 17 Reliability Report presents ongoing efforts using test-to-fail methodology to develop more comprehensive and advanced lifetime models, which is aimed at accurately projecting the reliability of GaN devices under more complex mission-specific operating conditions.

HERE ARE THE NEW ADDITIONS TO THE PHASE 17 RELIABILITY REPORT

The latest Phase 17 reliability report further expands the first-principles lifetime models to address more complex operating conditions, enabling more accurate lifetime projections for mission specific applications. Additionally, the latest version focuses on presenting the complex physics-based models in a variety of application-driven, user-friendly formats, allowing readers to quickly comprehend the concepts and apply them to practical use conditions with ease.

Section 4.1 presents an expanded gate lifetime model which now incorporates the effect of gate leakage current under various gate-source voltages and temperatures into the dominant impact ionization mechanism. Next, a duty cycle-based repetitive transient gate overvoltage rating at 7 V was developed and validated through the development of a repetitive inductive-switching gate overvoltage testing system, which accurately models the resonance-like transient gate overvoltage stress during applications.

Section 4.2 provides more testing validation to the repetitive transient drain overvoltage specification, which shows the excellent robustness of GaN devices under drain-source overvoltage conditions.

Section 4.3.5 is a new section that presents the latest development and measurements to quantify the pulsed current rating for the generation-6 and generation-5 GaN devices at various gate drive voltages and temperatures. The testing was also extended to more than 100 million pulses, after which minimal parametric shifts were observed. This work paves the way for GaN in applications that require transient high current pulses, such as light detection and ranging (lidar).

Section 4.4 presents the development of a comprehensive lifetime model for thermomechanical stress, applicable to both temperature cycling (TC) stress and power cycling (PC) stress. This revamped section further enhances the completeness of the TC lifetime modeling by incorporating die dimensions, bump shape, TC test conditions, ramp rate, and PCB properties and PCB thickness. Additionally, this section extends the thermomechanical discussions and lifetime modeling to power cycling for the first time. In PC, the temperature rise results from device self-heating during power-on, while the PCB temperature lags behind, creating a non-uniform thermal gradient from the device to the PCB. Section 4.4.4 discusses and models the critical parameters involved in PC stress, including cycling time, internal die dimensions and geometry within a PQFN package, and temperature variation between two extremes.

TABLE OF CONTENTS

SECTION 1. NEW FOCUS AND ADDITIONS OF THE PHASE 17 RELIABILITY REPORT

Compared to the previous Phase 16 reliability report [1], the latest version focuses on further expanding the previously developed physics-based lifetime models to encompass more complex operating conditions. This expansion includes the lifetime models for gate voltage stress, temperature cycling thermo-mechanical stress, and power cycling thermo-mechanical stress. The development of more comprehensive lifetime models leads to more accurate lifetime projections for mission specific operating conditions.

The second highlight of the Phase 17 reliability report is the translation of the complex lifetime models into a variety of user-friendly formats, allowing readers to easily apply the models to practical use conditions. Additionally, this effort allows us to update our transient overvoltage rating and pulsed current specifications in the datasheets, making them more application-oriented and competitive compared to other GaN manufacturers. The results also demonstrate the excellent robustness of EPC's GaN devices.

SECTION 2. DETERMINING WEAR-OUT MECHANISMS USING TEST-TO-FAIL METHODOLOGY

Standard qualification testing for semiconductors typically involves stressing devices at or near the limits specified in their datasheets for a prolonged period, or for a certain number of cycles. The goal of standard qualification testing is to have zero failures out of a relatively large group of parts tested.

This type of qualification testing is inadequate since it only reports parts that passed a very specific test condition. By testing parts to the point of failure, an understanding of the amount of margin between the datasheet limits can be developed, and more importantly, an understanding of the intrinsic failure mechanisms can be found. By knowing the intrinsic failure mechanisms, the root cause of failure, and the behavior of this mechanism over time, temperature, electrical or mechanical stress, the safe operating life of a product can be determined over a more general set of operating conditions (For an excellent description of test-to-fail methodology for testing semiconductor devices, see reference [2]).

As with all power transistors, the key stress conditions involve voltage, current, temperature, and humidity, as well as various mechanical stresses. There are, however, many ways of applying these stress-conditions. For example, voltage stress on a GaN transistor can be applied from the gate terminal to the source terminal (V_{GS}), as well as from the drain terminal to the source terminal (V_{DS}). For example, these stresses can be applied continuously as a DC bias, they can be cycled on-and-off, or they can be applied as high-speed pulses. Current stress can be applied as a continuous DC current, or as a pulsed current. Thermal stress can be applied continuously by operating devices at a predetermined temperature extreme for a period of time, or temperature can be cycled in a variety of ways.

By stressing devices with each of these conditions to the point of generating a significant number of failures, an understanding of the primary intrinsic failure mechanisms for the devices under test can be determined. To generate failures in a reasonable amount of time, the stress conditions typically need to significantly exceed the datasheet limits of the product. Care needs to be taken to make certain the excess stress condition does not induce a failure mechanism that would never be encountered during normal operation. To make certain that excess stress conditions did not cause the failure, the failed parts need to be carefully analyzed to determine the root cause of their failure. Only by verifying the root cause can a complete understanding of the behavior of a device under a wide range of stress conditions be developed. As the intrinsic failure modes in eGaN® devices are better understood, two facts have become clear; (1) eGaN devices are more robust than Si-based MOSFETs, and (2) silicon MOSFET intrinsic failure models do not generally apply when predicting eGaN device lifetime under extreme or long-term electrical stress conditions.

Table 2-1 lists in the left-hand column all the various stressors to which a transistor can be subjected during assembly or operation. Using the various test methods listed in the third column from the left, and taking devices to the point of failure, the intrinsic wear-out mechanisms can be discovered. The wear-out mechanisms confirmed as of this writing are shown in the column on the right.

Stressor	Device/ Package	Test Method	Intrinsic Failure Mechanism
Voltage	Device	HTGB	Dielectric failure (TDDB)
			Threshold shift
		HTRB	Threshold shift
			$R_{DS(on)}$ shift
		ESD	Dielectric rupture
Current	Device	DC Current (EM)	Electromigration
			Thermomigration
Current + Voltage (Power)	Device	SOA	Thermal Runaway
		Short Circuit	Thermal Runaway
Voltage Rising/Falling	Device	Hard-switching Reliability	$R_{DS(on)}$ shift
Current Rising/Falling	Device	Pulsed Current (Lidar reliability)	None found
Temperature	Package	HTS	None found
Humidity	Package	MSL1	None found
		H3TRB	None found
		AC	None found
		Solderability	Solder corrosion
		uHAST	Denrite Formation/Corrosion
Mechanical / Thermo-mechanical	Package	TC	Solder Fatigue
		IOL	Solder Fatigue
		Bending Force Test	Delamination
		Bending Force Test	Solder Strength
		Bending Force Test	Piezoelectric Effects
		Die shear	Solder Strength
		Package force	Film Cracking
Radiation	Device	Gamma Radiation	None found
		Neutron Radiation	None found
		Heavy Ion Bombardment	Crystal displacement damage and ionization damage

Table 2-1: Stress conditions and intrinsic wear-out mechanisms for GaN transistors

SECTION 3: USING TEST-TO-FAIL RESULTS TO PREDICT DEVICE LIFETIME IN A SYSTEM

When multiple failure mechanisms or stressors are involved, the total failure rate of a system, commonly known as Failure in Time (FIT), is the sum of the failure rates per failure mechanism [3,4] as shown below,

$$FIT_{Total} = FIT_1 + FIT_2 + \cdots + FIT_i \quad \text{Eq. 3-1}$$

where FIT is failure in time, which typically represents the number of failures in 10^9 (1 billion) device hours, and the subscript indicates the different failure mechanisms identified.

FIT is inversely proportional to mean time to failure (MTTF) as described by

$$FIT = \frac{10^9}{MTTF} \quad \text{Eq. 3-2}$$

Therefore, by plugging Equation 3-2 into Equation 3-1, the total MTTF can be described by Equation 3-3,

$$\frac{1}{MTTF_{Total}} = \frac{1}{MTTF_1} + \frac{1}{MTTF_2} + \cdots + \frac{1}{MTTF_i} \quad \text{Eq. 3-3}$$

The subscripts are assigned to the reliability stressors that are relevant to the application of interest. Based on Equation 2-3, it is noted that the smallest denominator yields the smallest MTTF and therefore dominates the overall lifetime. It is critical to understand which stressor is the limiting factor in reliability because the weakest link warrants the most consideration during design and operations.

In most applications, devices experience various stress conditions over the course of the entire mission lifespan, including a combination of different bias conditions and different temperature profiles. Each stress condition corresponds to a specific failure rate (failures per unit time), specified as FR_a, FR_b, …, FR_n. The respective duration of each stress condition is denoted as t_a, t_b, …, t_n. Assuming $t_{total} = t_a + t_b + ... + t_n$ is 10^9 hours, the FIT calculation of total number of failures is then generalized for specific reliability stress conditions as

$$FIT = FR_a \cdot t_a + FR_b \cdot t_b + \cdots + FR_n \cdot t_n \quad \text{Eq. 3-4}$$

The time-averaged failure rate FR can be calculated as

$$FR = FR_a \frac{t_a}{t_{total}} + FR_b \frac{t_b}{t_{total}} + \cdots + FR_n \frac{t_n}{t_{total}} \quad \text{Eq. 3-5}$$

which can be simplified by introducing fractional operation time,

$$n = \frac{t_n}{t_{total}} \quad \text{Eq. 3-6}$$

noted as a, b, …, n. The sum of a, b, …, n is 100% which is given in Equation 3-7.

$$a + b + \cdots + n = 100\% \quad \text{Eq. 3-7}$$

Now Equation 3-5 can be simplified to

$$FR = FR_a \cdot a + FR_b \cdot b + \cdots + FR_n \cdot n \quad \text{Eq. 3-8}$$

It is known that the failure rate under each sub-stress condition is inversely proportional to the device lifetime LT [4] when the same number of failures is generated. The relation is shown in Equation 3-9.

$$FR \propto \frac{1}{LT} \quad \text{Eq. 3-9}$$

Plugging Equation 3-9 into Equation 3-8 yields Equation 3-10.

$$\frac{1}{LT_{Total}} = \frac{a}{LT_a} + \frac{b}{LT_b} + \cdots + \frac{n}{LT_n} \quad \text{Eq. 3-10}$$

where LT_{Total} is the total projected lifetime and LT_i is the projected lifetime for each stress condition.

Equation 3-10 captures how a mission profile consisting of more than one stress condition results in a system lifetime. The fractional operation time (a, b, …, n) in the numerators account for the times spent in harsh, moderate, and mild stress conditions.

SECTION 4: WEAR-OUT MECHANISMS

4.1. Gate Wear-Out

4.1.1. Introduction to the Reliability of Schottky-type pGaN Gates

Schottky-type pGaN gates are the most widely used gate structure for commercial enhancement-mode GaN HEMTs that are currently in volume production. A Schottky-type pGaN gate typically consists of gate electrode made of titanium nitride (TiN) and a pGaN gate layer that is doped with Mg. Due to the significant structural differences in gate construction between GaN HEMTs and Si-based MOSFETs, the stability and robustness of pGaN gates are of great interest to the users.

In this section, after understanding the fundamental gate wear-out mechanism through test-to-fail, a physics-based gate lifetime model was developed from first principles. The model predicts a failure rate of less than one part per million (1-ppm) if the gate bias is kept below 6 $V_{GS,Max}$ throughout the entire mission lifespan of ~25 years. The projected result is also consistent with EPC's field experience.

In this new phase 17 reliability report, the gate lifetime model has been further expanded to include the voltage and temperature dependence of the electron injection current density by analyzing the gate leakage current conduction mechanisms. This inclusion enables the accurate modeling of the activation energy of mean-time-to-fail (MTTF) at various temperatures.

Another common reliability question regarding Schottky-type pGaN gates is the transient overvoltage capability and robustness, due to the relatively small margin between the recommended gate drive voltage (~5 V) and the datasheet maximum specification ($V_{GS,Max}$ = 6 V). The latest phase 17 reliability report developed a 7 V repetitive transient overvoltage rating with 1% duty cycle factor,

which was later validated through the development of a repetitive inductive switching testing circuit.

4.1.2. Development of a Comprehensive Gate Reliability Lifetime Model

To understand the gate wear-out mechanisms, accelerated time-dependent reliability testing was conducted on various EPC's GaN HEMTs at various voltages and temperatures. Failure analysis revealed that the breakdown of the silicon nitride dielectric layer, located between the gate corner and metal field plate, is primarily responsible for the pGaN gate failures, as shown in Figure 4-1.

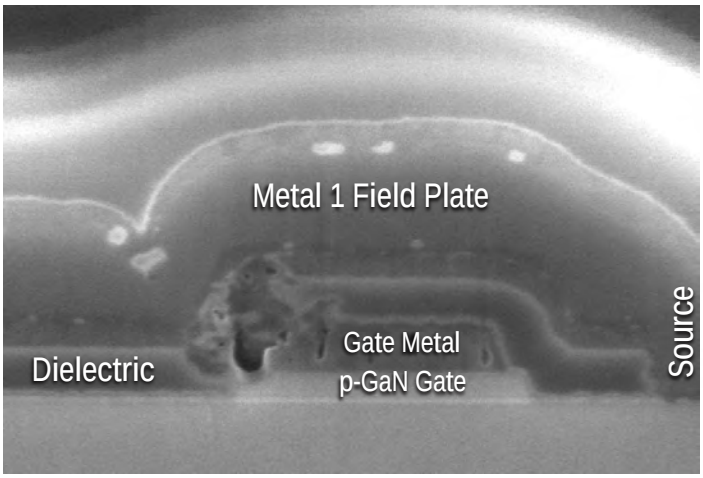


Figure 4-1. Scanning electron microscopy (SEM) image of a gate failure. Dielectric breakdown is observed between the gate metal and the field plate metal.

Impact ionization was identified as the main wear-out mechanism responsible for the silicon nitride dielectric breakdown failure mode [5]. A four-step process was developed to explain the failure mode shown in Figure 4-1. The electron injection from the 2-dimensional electron gas (2DEG) and the subsequent acceleration within the pGaN gate layer is the first step. When the pGaN gate is subjected to a high forward gate bias (V_{GS}), the 2DEG electrons fully populate the channel and may spill over the "bending" conduction band of the AlGaN barrier layer. Subsequently, the injected electrons are accelerated within the depleted pGaN gate layer under high forward V_{GS}, gaining significant kinetic energy [5]. When the energetic moving electrons are stopped by the TiN gate metal/pGaN interface, the resulting bombardment causes impact ionization and triggers electron-hole multiplication, which has been confirmed by luminescence measurements [6]. Thus, Impact ionization and electron-hole multiplication at the TiN/pGaN interface constitute the second step. The third step involves hole accumulation within the silicon nitride dielectric layer. The positively charged holes generated by impact ionization move away from the gate electrode (under + V_{GS}) towards the metal field plate that is at ground potential during gate stress. Consequently, the holes become trapped in the silicon nitride dielectric layer, leading to an increasing positive charge density as the gate stress continues. Finally, when the trapped hole density exceeds the critical field of the silicon nitride dielectric layer, dielectric breakdown occurs, which explains the failure mode as shown in Figure 4-1. Based on the four-step impact ionization failure process, a physics-based gate lifetime model was developed from first principles.

The MTTF is modeled by estimating when the trapped hole charges reach the critical charge density (Q_C) of the silicon nitride dielectric layer, as defined by Eq. 4-1:

$$MTTF = \frac{Q_C}{G} \quad \text{Eq. 4-1}$$

where G is the electron-hole generation rate ($s^{-1}cm^{-3}$) that is denoted by Eq. 4-2. It is noted that holes are the primary cause responsible for the dielectric breakdown.

$$G = \alpha_n \frac{J_n}{q} \quad \text{Eq. 4-2}$$

where Jn is the electron current density (A/cm^2) that is directly proportional to the gate leakage current under forward gate bias, q is the elementary charge (coulomb = A-s), and αn is the electron impact ionization coefficients (cm^{-1}), which is defined by the Chynoweth model in Eq. 4-3 [7].

$$\alpha_n = a_n e^{-(\frac{b_n}{E})^m} \quad \text{Eq. 4-3}$$

E is the vertical electric field driven by gate bias and m is an exponent that is typically ranging from 1 to 2; a_n and b_n are temperature dependent impact ionization coefficients that can be described by the Okuto-Crowell model [8], which are further defined by Eq. 4-4 and Eq. 4-5 [9].

$$a_n = a_{n,0}(1 + c\Delta T) \quad \text{Eq. 4-4}$$

$$b_n = b_{n,0}(1 + d\Delta T) \quad \text{Eq. 4-5}$$

where ΔT is the temperature difference compared to 298 K in Kelvin unit [17-19]. $a_{n,0} = 2.77 \times 10^8\ cm^{-1}$, $b_{n,0} = 3.20 \times 10^7$ V/cm, $c = 3.09 \times 10\text{-}3\ K^{-1}$, $d = 5.03 \times 10^{-4}\ K^{-1}$ are the fitting parameters of impact ionization coefficients by following the Okuto-Crowell model [8]. By combining Eq. 4-1 – Eq. 4-5, the MTTF becomes Eq. 4-6:

$$MTTF = \frac{qQ_C}{J_n a_{n.0}(1+c\Delta T)} e^{\left(\frac{b_{n.0}(1+d\Delta T)}{E}\right)^m} \quad \text{Eq. 4-6}$$

First, time-dependent gate reliability testing was conducted on EPC2212 under four different gate biases: 8 V, 8.5 V, 9 V and 9.5 V at room temperature of 25°C. Therefore, ΔT is equal to 0, leading to a simplified MTTF expression as shown in Eq. 4-7.

$$MTTF = \frac{qQ_C}{J_n a_{n,0}} e^{\left(\frac{b_{n,0}}{E}\right)^m} \quad \text{Eq. 4-7}$$

where m is 1.9, $a_{n,0} = 2.77 \times 10^8\ cm^{-1}$, and $b_{n,0} = 3.20 \times 10^7$ V/cm.

Figure 4-2 shows that the gate lifetime equation of Eq. 4-7 provides a good fit to the measured MTTF at various gate biases. Additionally, less than 1-ppm (part per million) failure rate is predicted if the gate bias is kept at or below the maximum gate rated voltage of 6 V for 25 years.

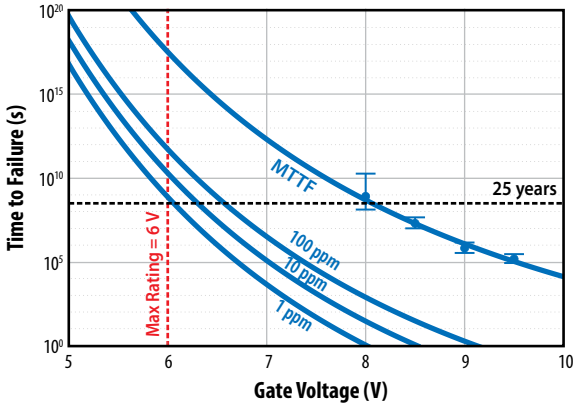


Figure 4-2: EPC2212 MTTF vs. V_{GS} at 25°C (and error bars) are shown for four different voltage legs. The solid line corresponds to the impact ionization lifetime model. Extrapolations of time to failure for 100 ppm, 10 ppm, and 1 ppm are shown as well.

Next, time-dependent gate reliability was carried out at various temperatures with a fixed gate bias of 9.5 V on EPC2057. The Weibull distribution plot at three different temperatures (-25°C, 25°C and 125°C) is shown in Figure 4-3. When the temperature increases from -25°C to 25°C, the gate lifetime increases, suggesting a negative activation energy (E_a). However, as the temperature continues rising further to 125°C, the gate lifetime decreases, indicating a positive E_a. This suggests that two competing effects are likely responsible for the pGaN gate breakdown failures.

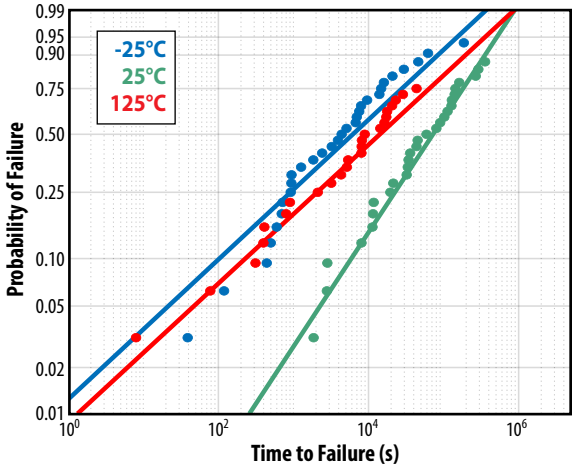
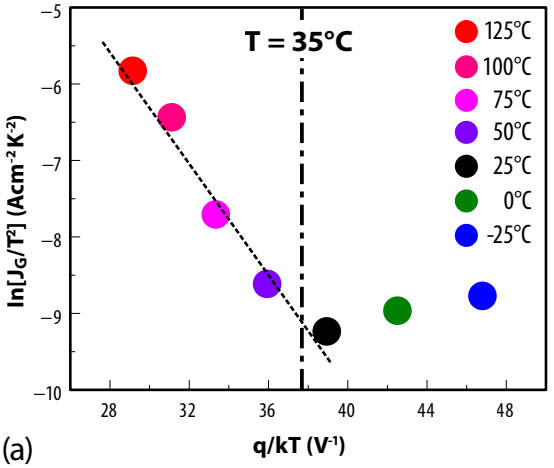


Figure 4-3: Weibull distribution plots of EPC2057 under three different temperatures: -25°C, 25°C and 125°C with a fixed gate bias of 9.5 V.

To develop a comprehensive gate lifetime model, the voltage and temperature dependence of J_n must be further investigated, where J_n is directly proportional to the forward gate leakage current (I_G). Therefore, the gate leakage current in EPC2057 was measured at different temperatures and voltages, with the gate I-V results reported in [10]. A significant temperature acceleration of I_G is observed at higher temperatures, suggesting that thermionic emission (TE) is proposed as the dominant conduction mechanism, which can be modeled by Richardson's law [11], as shown in Eq. 4-8.

$$J_{TE} = AT^2 e^{\left(\frac{-\varphi_B}{kT}\right)} \quad \text{Eq. 4-8}$$

where A is the Richardson's constant, k is the Boltzmann constant, and φ_B is the barrier height for electrons to overcome the AlGaN/GaN heterojunction. φ_B is calculated to be 0.45 eV at 9.5 V_{GS} based on the slope of the fit line shown in Figure 4-4 (a).



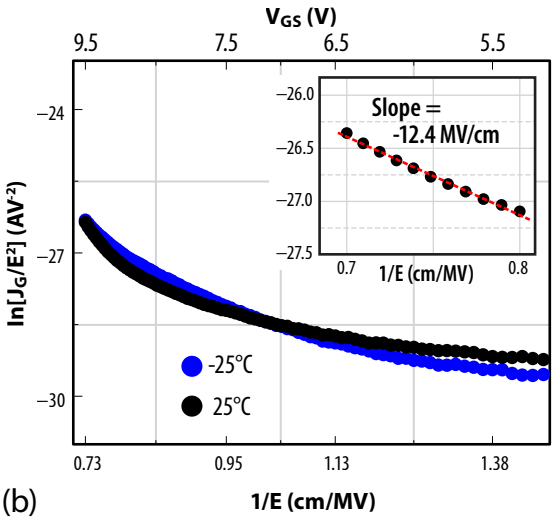


Figure 4-4: (a) Richardson plot from -25°C to 125°C with 9.5 V_{GS}; (b) FN plot at -25°C and 25°C, where the inset shows the linear fits from 9 V to 9.5 V at 25°C.

After combining Eq. 4-6 and Eq. 4-8, the MTTF at higher temperatures can be written as Eq. 4-9

$$MTTF = \frac{qQ_c}{AT^2 a_{n,0}(1+c\Delta T)} e^{[(\frac{b_{n,0}(1+d\Delta T)}{E})^m + \frac{\varphi_B}{kT}]} \quad \text{Eq. 4-9}$$

In Figure 4-4 (a), the Richardson plot produces a straight line fit from 50°C to 125°C, which confirms that TE is the dominant conduction mechanism. However, the data points from 25°C to -25°C deviate from the fit line, indicating that TE is no longer the primary conduction mechanism responsible for I_G. Thus, 35°C is projected to be the threshold temperature at which the dominant I_G conduction mechanism transitions from TE to other mechanisms. When plotting $\ln(J_G/E^2)$ against 1/E as shown in Figure 4-4 (b) for 25°C and -25°C, insignificant dispersion is observed when the V_{GS} is greater than 9 V, suggesting that Fowler-Nordheim (FN) tunneling is the dominant conduction mechanism [12]. The FN tunneling can be modeled by Eq. 4-9.

$$J_{FN} = \frac{q^2(m_e/m^*)}{8\pi h\varphi_{eff}} E^2 e^{\frac{-8\pi\sqrt{2m^*(q\varphi_{eff})^3}}{3hqE}} \quad \text{Eq. 4-10}$$

where φ_{eff} is the effective barrier height, h is the Planck constant, and m* is the electron effective mass. It is widely reported that the φ_{eff} for FN tunneling at low temperatures is found consistent with the φ_B from TE at high temperatures [13]. Hence, 0.45 eV is adopted for φeff in FN tunneling. Based on the FN slope of -12.4 MV/cm, m* is estimated to be ~0.36 m_e, matching the commonly reported m* of 0.4 m_e in AlGaN [14].

Combining Eq. 4-6 and Eq. 4-10 yields the MTTF at low temperatures and higher gate biases, as shown in Eq. 4-11.

$$MTTF = \frac{8\pi h\varphi_{eff}Q_c}{q(m_e/m^*)E^2 a_{n\,0}(1+c\Delta T)} e^{[(\frac{b_{n,0}(1+d\Delta T)}{E})^m + \frac{8\pi\sqrt{2m^*(q\varphi_{eff})^3}}{3hqE}]} \quad \text{Eq. 4-11}$$

After further expanding the current density term (J_n) as shown Eq. 4-8 and 4-10, now a comprehensive gate lifetime model can be developed as shown in Eq. 4-9 and Eq. 4-11. When T < 35°C and V_{GS} = 9.5 V, J_n is dominated by FN tunneling conduction mechanism that shows minimal temperature dependence. Hence, the gate lifetime is dominated by the impact ionization coefficient (αn). The injected hot electrons experience less lattice scattering [7,9], leading to more energetic bombardment, a higher hole generation rate (G) and shorter MTTF. Figure 4-5 shows that the E_a is estimated to be -0.19 eV from -25°C to 25°C, calculated by Eq. (4-12). At higher temperatures, αn decreases due to increased phonon scattering, which prevents the injected electrons from gaining sufficient energy and slows down the gate wearout. However, I_G increases exponentially at higher temperatures, leading to orders of magnitude more electrons being injected, which accelerates gate wearout. Figure 4-5 shows that from 25°C to 125°C, the increase in hot electrons overwhelms the decrease in α_n. As illustrated in Figure 4-5, a positive E_a of 0.1 eV is measured between 25°C and 125°C with Eq. 4-12.

$$MTTF \propto Exp(E_a/kT) \quad \text{Eq. 4-12}$$

For more detailed analysis of the comprehensive gate lifetime modeling, please refer to this journal publication on IEEE Electron Device Letters by EPC [10].

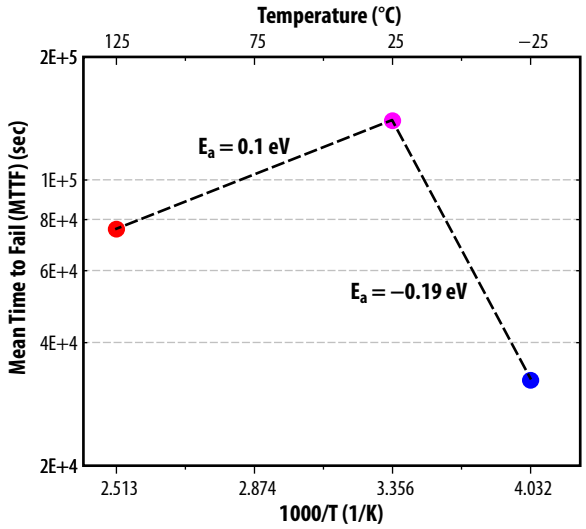


Figure 4-5: Experimentally measured E_a at different temperatures with V_{GS} = 9.5 V for EPC2057.

4.1.3. Development of a Repetitive Transient Gate Over-voltage Specification

Gate overvoltage spikes during device turn-on transients are commonly observed in GaN HEMTs under high-frequency, fast-switching conversion applications [15,16]. The magnitude of the gate overvoltage transients is primarily governed by the gate-loop inductance and the slew rate (V_{GS}/dt) which both are closely related to circuit design and PCB layout [17].

Figure 4-2 projects nearly zero failure rate if the gate drive voltage does not exceed the maximum gate voltage rating of 6 V. It is consistent with EPC's field experience, where no gate failures have been identified despite very demanding applications in automotive, satellites, and advanced enterprise servers. However, the reliability and robustness under repetitive gate overvoltage stress are frequently inquired by the users. Therefore, there is a strong demand to develop a repetitive transient gate overvoltage specification supported by reliability data. Figure 4-2 also predicts that when the gates are subjected to 7 V continuous gate bias, the estimated gate lifetime is 3.3 x 10^6 seconds with 100-ppm failure rate. When comparing to a typical mission lifespan of 10 years (3.1 x 10^8 seconds), this corresponds to slightly more than 1% of the total lifespan. The 1% can then be translated to a duty cycle factor (DC_{Factor}) that occurs in every switching cycle as shown in Figure 4-6.

In real-world applications, the transient gate overvoltage profile can be illustrated by the simplified waveform shown in Figure 4-6, where T_S represents the switching period, which is the inverse of the switching

frequency, and T_O is the duration of the transient overvoltage ringing. Thus, the DC_{Factor} can be defined as the ratio between T_O and T_S, suggesting that GaN HEMTs should be able to sustain a repetitive 7 V_{GS} overvoltage spike lasting 1% of each switching period, while maintaining a low failure rate.

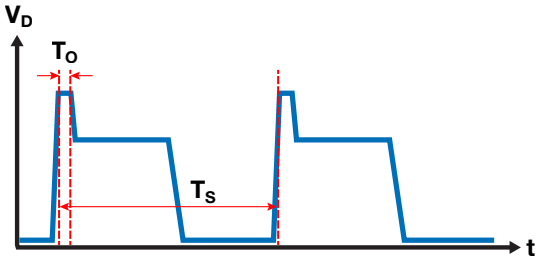


Figure 4-6: Illustration of the 1% DC_{Factor} overvoltage specification, which is defined by the ratio between T_O (overvoltage duration) and T_S (switching period).

To validate the 1% DC_{Factor} overvoltage specification, an inductive switching test system was developed to emulate the gate overvoltage ringing phenomenon observed in switching applications. Figure 4-7 (a) shows the schematic of the inductive switching test circuit. The circuit operation can be divided into two phases: Phase 1, the charging phase and Phase 2, the transient phase. During Phase 1, FET Q in Figure 4-7 (a) is turned on for a specified time interval (t), during which the input voltage (V_{IN}) charges the inductor (L), causing the inductor current to rise, as shown in Figure 4-7 (b). Next, when the FET Q is turned off, the inductor (L) and parasitic capacitance (C) generate an LC resonance, leading to a half-sinusoidal transient overvoltage spike with a peak voltage of 7 V, as shown in Figure 4-7 (b). The time interval (t) is adjusted to ensure that the pulse width of V_{GS} > 6 $V_{GS,Max}$ stays consistently at ~70 ns for all parts. The testing switching frequency is approximately 3 MHz.

Four different GaN HEMT products and three parts per product with a drain-source (V_{DS}) rating from 50 V to 200 V were tested with a peak V_{GS} of 7 V to a trillion pulses at 25°C. Figure 7 shows the evolution of threshold voltage (V_{TH}) and on-resistance ($R_{DS(on)}$) of a representative device from each product. Device characterization was conducted prior to testing and after reaching a trillion cycles. As shown in Figure 4, the post-stress measurements are well below the datasheet limit of each product.

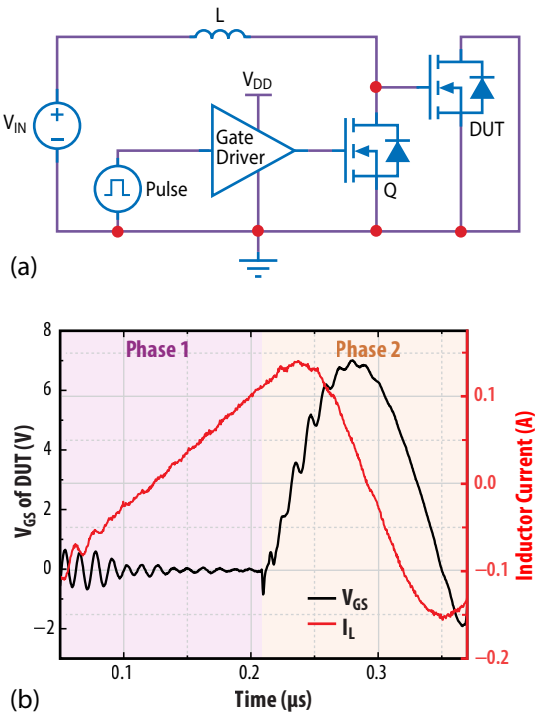


Figure 4-7: (a) a simplified schematic of the inductive gate switching test system; (b) the measured inductor current and V_{GS} waveforms with a peak voltage of 7 V.

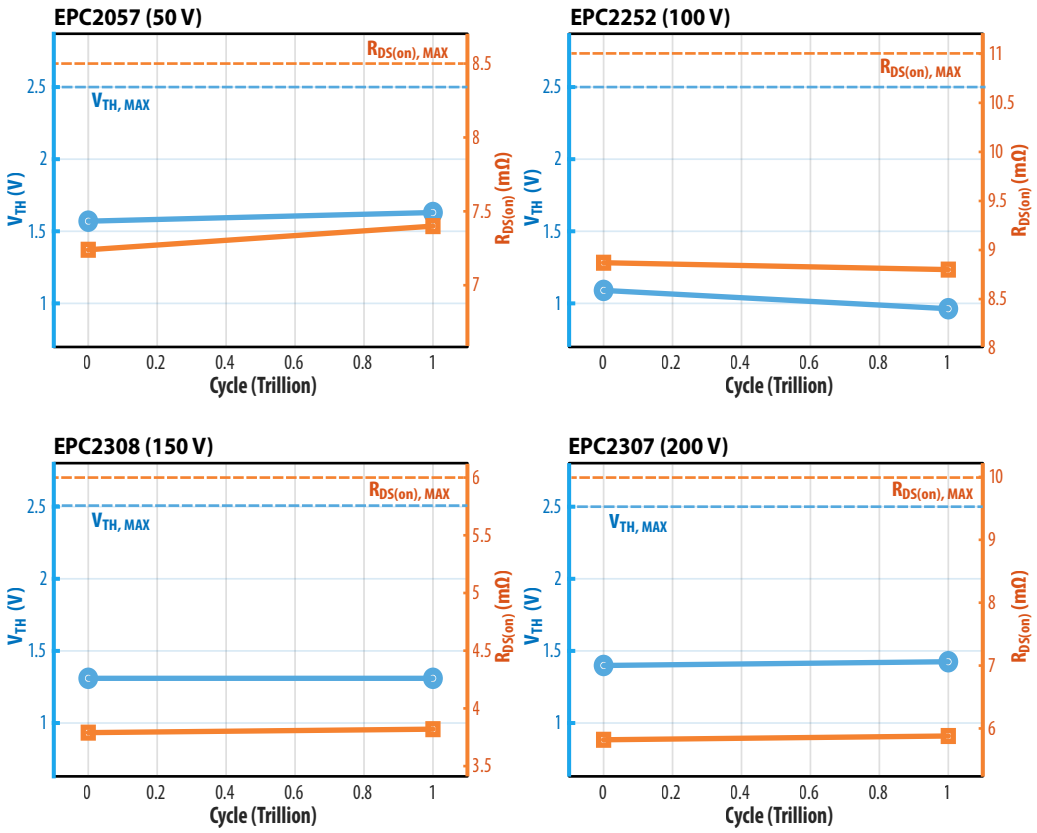


Figure 4-8. Parametric comparison of pre- and post- stress, 1-trillion gate overvoltage spikes with a peak voltage of 7 V of four representative GaN HEMT products, including EPC2057, EPC2252, EPC2308, and EPC2307.

Additionally, two representative products (EPC2057 and EPC2307) were subjected to another trillion pulses of stress with a peak V_{GS} of 7 V, while the junction temperature was maintained at 125°C. Figure 4-9 shows the static parameter measurements after an additional trillion cycles of stress at 125°C, where no significant shift was observed.

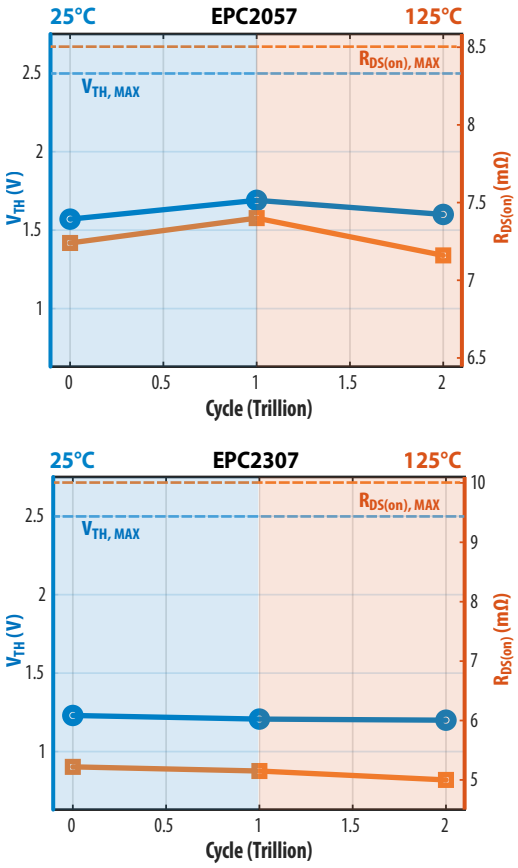


Figure 4-9. On-resistance ($R_{DS(on)}$) and threshold voltage (V_{TH}) parametric comparison of pre- and post-stress, 2-trillion pulses with peak voltage of 7 V at 25°C (blue shaded) and 125°C forced heating (orange shaded) of EPC2057 and EPC2307.

A total of 12 GaN HEMTs have been subjected to a total of 15 trillion pulses of stress with a peak gate voltage of 7 V at the time of writing this report. Since each device was tested with a consistent time interval of approximately 70 ns with V_{GS} > 6 V, the total stress time is calculated to be approximately 1.1 x 10^6 seconds, which is one third of the projected gate lifetime at 7 V static gate bias with 100-ppm failure rate (3.3 x 10^6 seconds). Since no observable parameter shift has been detected, this suggests that there is still a significant margin in lifetime before the GaN HEMTs show any measurable parametric degradation. More testing is underway to further validate the applicability of the proposed 1% DC_{Factor} specification. However, the inductive gate overvoltage switching test results to date support a repetitive transient gate overvoltage rating of 7 V with a 1% DC_{Factor}.

To demonstrate how to implement the 1% DC_{Factor} overvoltage specification, an example is provided. If a converter operates at 1 MHz switching frequency (T_S = 1 µs), a repetitive overvoltage spike occurs during the gate turn-on transients due to unoptimized gate loop inductance. The spike has a peak V_{GS} of 7 V with a time interval of 8 ns above 6 V_{GS}. Dividing 8 ns by the 1 µs of T_S yields 0.8%, which is below the 1% DC_{Factor}. Therefore, a failure rate much lower than 100 ppm is expected after 10 years of continuous operation.

4.2. Drain Wear-Out

4.2.1. Introduction to Drain Wear-Out Mechanisms

Dynamic on-resistance ($R_{DS(on)}$) is one of the most common reliability concerns for GaN HEMTs when subjected to high drain-source bias stress. Dynamic $R_{DS(on)}$ refers to the condition in which the on-resistance of the GaN HEMTs increases when the device is exposed to high drain-source voltage (V_{DS}).

In this section, a similar test-to-fail method is used to investigate drain-related wear-out mechanisms. After understanding the underlying mechanisms responsible for dynamic $R_{DS(on)}$, a comprehensive physics-based drain lifetime model was developed from first principles to project dynamic $R_{DS(on)}$ shifts with respect to various parameters, including voltage, temperature, frequency, and current.

GaN HEMTs are increasingly deployed in advanced applications, featuring high switching frequencies and fast slew rates. Thus, reliability and robustness under repetitive transient drain overvoltage stress have become another frequently asked reliability question by users. Later in this section, a similar duty cycle-based repetitive drain overvoltage specification was developed by using a resistive hard-switching testing circuit, which was subsequently validated through the development of an inductive switching test circuit.

4.2.2. Development of Physics-Based Lifetime Models for Dynamic $R_{DS(on)}$

As discussed in the previous reliability reports [1], the dominant mechanism responsible for the dynamic $R_{DS(on)}$ failure mode is electron trapping at or near high electric field regions, leading to the depletion of 2DEG electrons within the drift region. Figure 4-10 shows a magnified image of an EPC2016C GaN HEMT displaying thermal emissions in the 1–2 µm optical range. These emissions observed in such wavelength range are consistent with hot electron mechanism. After aligning the emissions with the device layout, it was found that these hot electron emissions occur in areas where the highest electric fields are present under high drain-source bias. This critical finding has led to the development of the next generation GaN HEMTs in which the peak electric fields are carefully managed to minimize dynamic $R_{DS(on)}$. Therefore, the latest generation GaN HEMTs exhibit nearly no dynamic $R_{DS(on)}$.

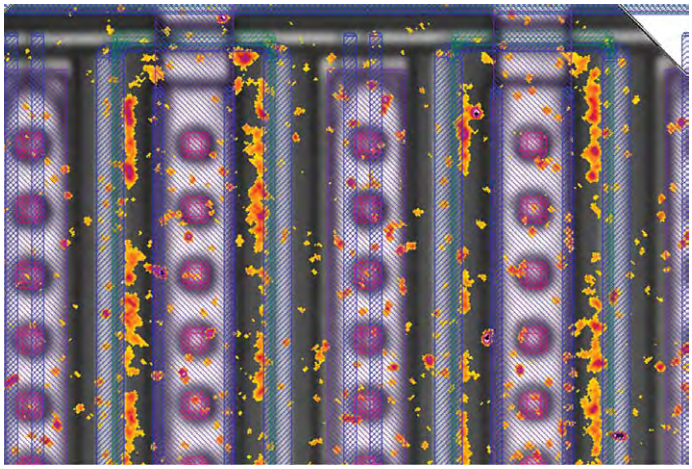


Figure 4-10: A magnified image of an EPC2016C GaN transistor showing light emission in the 1–2 µm wavelength short-wave infrared light range (SWIR) that is consistent with hot electrons. The SWIR emission (red-orange) has been overlaid on a regular (visible wavelength) microscope image and a semi-transparent image of the design photomask (purple).

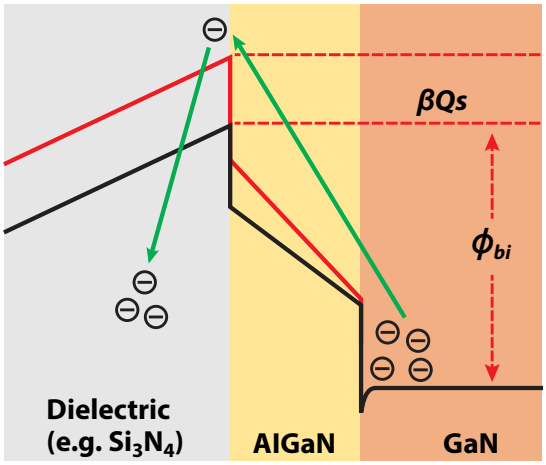


Figure 4-11: Illustration of the self-limiting trapping process, where the barrier height is enhanced after the most energetic electrons are trapped. The dynamic barrier change is be quantified as $\beta \times Q_S$, where β is a geometric factor that correlates the dynamic barrier height increase with respect to the trapped charges (Q_S).

After understanding the fundamental wearout mechanism responsible for dynamic $R_{DS(on)}$, a comprehensive lifetime model was developed to describe the rise in dynamic $R_{DS(on)}$ of GaN HEMTs. This model was also derived from first principles under hard-switching test conditions. The model is predicated on the assumption that hot electrons are injected over a surface potential into the conduction band of the dielectric layer (e.g. Si_3N_4), where the electric field is highest. Figure 4-11 illustrates the band structure at the interfaces of GaN layer/AlGaN barrier layer/Si_3N_4 dielectric layer. After the more energetic electrons overcome the barrier and become trapped in the dielectric layer, those trapped charges (Q_S) exert an additional electrostatic screening force against the electrons in the 2DEG, causing a dynamic barrier height increase. Further barrier height enhancement hinders other energetic 2DEG electrons from getting trapped, which leads to a self-limiting trapping process. Since these hot electrons are created during the hard-switching transitions, the transient combination of high injection current and high fields leads to a hot carrier energy distribution with long tails in the high energy regime.

This self-limiting electron trapping rate $\frac{dQ_S}{dt}$ can be modeled by the integral of the electron density distribution function (*f(E)*) bounded by the energy barrier $\phi_{bi} + \beta Q_S$ to infinity where virtually no electrons can overcome the energy barrier, and the trapping process ultimately stopped, as shown in Eq. 4-13.

$$\frac{dQ_S}{dt} = A\int_{\Phi_{bi}+\beta Q_S}^{\infty} f(E)dE \qquad \text{Eq. 4-13}$$

where the electron density distribution, *f(E)*, is exponentially dependent on electron energy (E), as shown in Eq. 4-14 [16,18].

$$f(E)dE \propto Ee^{-E/qF\lambda}dE \qquad \text{Eq. 4-14}$$

where *f* is electric field, q is electron charge, and λ is electron mean free path.

Therefore, the Q_S is solved and shown in Eq. 4-15.

$$\lambda = v_{th}\tau_{LO} \propto A\sqrt{kT}exp\left(\frac{\hbar\omega_{LO}}{kT}\right) \qquad \text{Eq. 4-15}$$

Under typical operation conditions, where the applied V_{DS} does not exceed 120% of the $V_{DS,Max}$, the Q_S is expected to be significantly less than the built-in piezoelectric charges in the 2DEG, Q_P [2,3]. Additionally, another assumption is that once the electrons are trapped, they are trapped permanently (no de-trapping). Therefore, the final expression to define the dynamic $R_{DS(on)}$ shift, $\Delta R_{DS(on)}/R_0$ is shown in Eq. 4-16.

$$\frac{\Delta R}{R} = \frac{R(t) - R(0)}{R(0)} \approx a + bF\exp\left(\frac{\hbar\omega_{LO}}{kT}\right)\sqrt{T}\log(t) \qquad \text{Eq. 4-16}$$

where V_{DS} is the drain-source voltage, T is device junction temperature in Kelvin unit, t is testing time in minutes. Other parameters in the mathematical model were fitted to the measured results across a range of drain voltages and temperatures, where a is a unitless fitting parameter, b equals 2.0E-5 ($K^{-1/2}$), $\hbar\omega_{L0}$ is 92 meV, corresponding to the LO phonon energy level scattered by the hot electrons, VFD is 100 V for generation-5 (Gen5) 100 V products only, and α equals 10 V.

Therefore, dynamic $R_{DS(on)}$ shift can be modeled by a linear relation with respect to logarithmic of test time (log-t) under hard-switching conditions. Figure 4-12 shows the voltage and temperature dependence of dynamic $R_{DS(on)}$ for a fifth-generation EPC2045 GaN HEMT with a maximum drain-source voltage rating of 100 V ($V_{DS,Max}$ = 100 V). The results showed that the measurement data points followed the logarithmic-time lifetime projection, validating the effectiveness of the lifetime model in Eq. 4-16.

On the top graph of Figure 4-12, the EPC2045 devices were tested at 25°C with the applied drain-source voltage ranging from 60 V to 120 V. The results show that the dynamic $R_{DS(on)}$ increases as a function of drain-source voltage (V_{DS}). As the V_{DS} increases, the peak electric field increases, which accelerates the hot electron trapping effect, leading to more significant dynamic $R_{DS(on)}$ rise over time. The graph on the bottom shows the time evolution of $R_{DS(on)}$ when biased at 120 V across three different temperatures: 25°C, 75°C and 125°C. The counter-intuitive result shows that dynamic $R_{DS(on)}$ effect becomes more prominent at lower temperatures than at higher temperatures, which is consistent with hot-carrier injection theory. At lower temperatures, these energetic electrons can travel further between scattering events from the LO-phonon, gaining greater kinetic energies under a given electric field. When the hot electrons are accelerated to higher energies, they can reach deeper layers in which charge trapping becomes more likely. This finding also suggests that traditional testing methods, such as high temperature reverse bias (HTRB), where a device is tested at maximum drain-source voltage and temperature for a long duration, may not be enough to determine the reliability of a device.

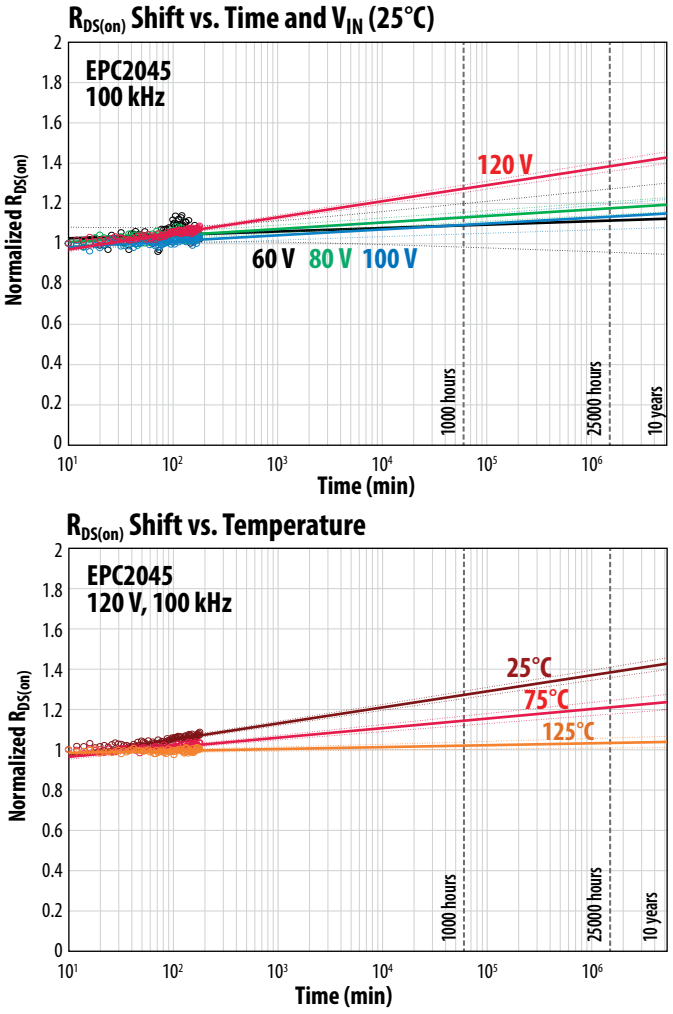


Figure 4-12: The $R_{DS(on)}$ of a fifth generation EPC2045 GaN transistor over time at various voltage stress levels and temperatures. On the top, the devices were tested at 25°C with voltages from 60 V to 120 V. The graph on the bottom shows the evolution of $R_{DS(on)}$ at 120 V at various temperatures.

The model allows users to project long-term $R_{DS(on)}$ growth as a function of four key input variables: drain voltage, temperature, switching frequency, and switching current with the following observations.

- $R_{DS(on)}$ growth with time
- The slope of $R_{DS(on)}$ over time has a negative temperature coefficient (i.e. lower slope at higher temperature)
- Switching frequency does not affect the slope, but causes a small vertical offset
- Switching current does not affect the slope
- Negligible difference between inductive and resistive hard switching.

4.2.3. Impact of Higher Drain-Source Voltage Stress

In the case where the amount of trapped charge approaches the number of electrons available in the 2DEG (the surface trapped charges (Q_S) approaches the built-in 2DEG piezoelectric charge (Q_P), the simplifying assumption used to develop Equation 4-16 is no longer valid. This situation could occur when devices are taken to voltages well above their design limits. Figure 4-13 shows results for EPC2045 devices tested up to 150 V at 75°C and 125°C. Note how the straight-line extrapolation that would occur with a simple log(time) dependence is no longer applicable. By removing the simplified assumption that only a small fraction of Q_P is trapped and transform into Q_S, the result shown in Eq. 4-17 is obtained. Calculating Eq. 4-17 using the expanded list of parameters yields the solid fit lines in Figure 4-13, providing further evidence of the validity and applicability of this physics-based model.

$$\frac{\Delta R}{R} = a_1\left[\frac{a_2\Psi \log\left(1+a_3 t/\Psi\right)}{1-a_2\Psi \log\left(1+a_3 t/\Psi\right)}\right] \qquad \text{Eq. 4-17}$$

where:

$$a_1 \equiv \frac{C}{Q_P} \qquad a_2 \equiv \frac{1}{Q_P} \qquad a_3 \equiv B$$

with the following expanded list of parameters:

a_1 = 0.6 (unitless)
a_2 = b/a_1 (where b = 2.0E-5 $K^{-1/2}$ from [19])
a_3 = 1000 ($K^{1/2}$ min^{-1})
b = 2.0E-5 ($K^{-1/2}$)
$\hbar\omega_{L0}$ = 92 meV
V_{FD} = 100 V (appropriate for Gen5 100 V products only)
α = 10 (V)
T = Device temperature (K)
t = Time (min)

Figure 4-14 compares this model to measurements of 200 V devices. On the left is the normalized $R_{DS(on)}$ for the fifth-generation, 200 V rated EPC2215 at three voltages. The highest voltage, 280 V, is 40% above the maximum rating. On the right are measurements compared with the model at two different temperatures and the maximum rated voltage.

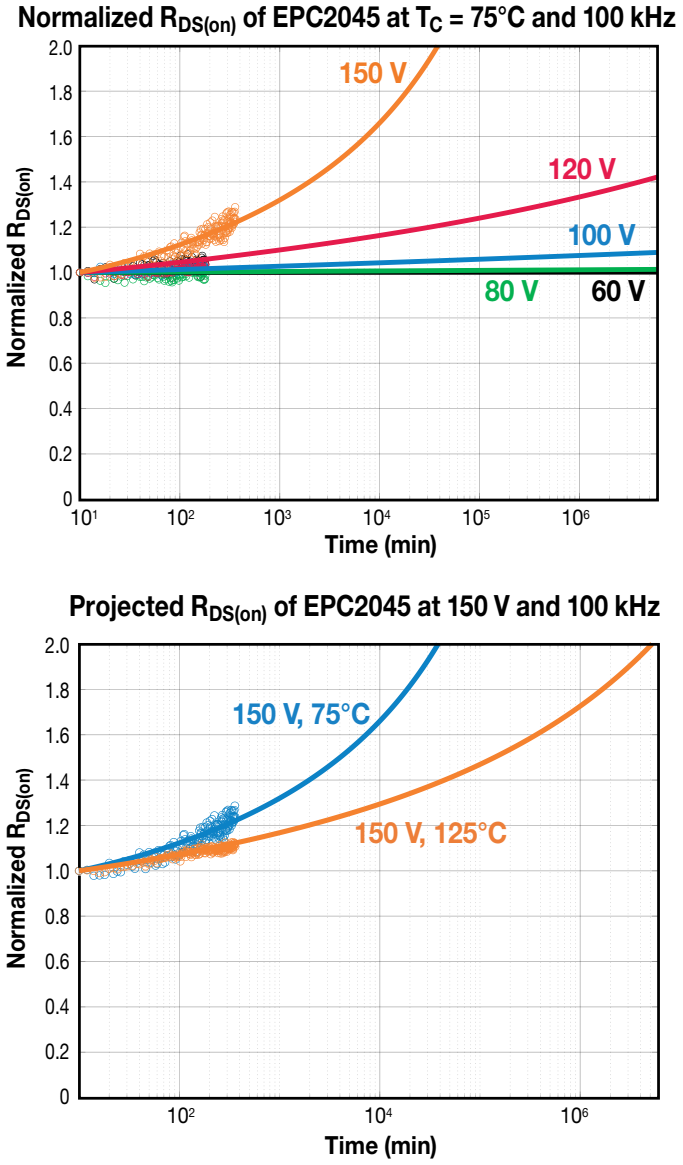


Figure 4-14: (Left) 200 V EPC2215 normalized $R_{DS(on)}$ at three voltages. Note that 280 V is 40% above the maximum rated voltage. (Right) EPC2215 at 75°C and 125°C and 200 V. The solid lines are the model results using variables for 200 V devices, and the dots are actual measurements.

4.2.4. Development of a Repetitive Transient Drain Overvoltage Specification

Transient drain voltage overshoot is commonly observed in GaN-based converters due to high slew rate and fast switching applications. A survey of transient overvoltage specification from a suite of GaN suppliers was conducted by JEDEC JC-70 committee and presented in JEP186 [20]. Most of the transient overvoltage specifications describe it as a device robustness indicator. In addition, many of them consider drain voltage overshoot as a single rare event or atypical occurrence. Hence, it is challenging for application engineers to effectively implement these specifications into their designs. Therefore, an application driven, and user-friendly repetitive transient off-state drain overvoltage specification on datasheets is important for the general adoption of GaN technology because of the absence of avalanche mechanisms in GaN HEMTs.

A resistive hard-switching test system [1,15] was employed to study dynamic $R_{DS(on)}$ shift under cumulative drain overshoot stress, where this system operates at 100 kHz, 85% of the time reverse-biasing the GaN device under test (DUT) at the specified off-state drain voltage. When determining time of failure, 20% of $R_{DS(on)}$ shift compared to the initial $R_{DS(on)}$ value after a projected 25 years of stress is used as the failure criteria. Eq. 4-14 is used to extrapolate the time-of-failure when the in-situ monitored $R_{DS(on)}$ shifts more than 20% to its initial value (R_0). This approach is more stringent than the typical datasheet maximum $R_{DS(on)}$ limit.

A suite of 100 V fifth generation GaN products were tested by the resistive hard switching test circuit at 120% of $V_{DS,Max}$ and 75°C junction temperature, a common mission temperature. EPC2045, the first generation-5 100 V drain-source rated GaN product, was subjected to testing under such accelerated hard-switching conditions. Figure 4-15 shows the testing results, where the DUT is projected to exceed the 20% $R_{DS(on)}$ shift limit at approximately 2 x 10^5 minutes by considering a 90% upper bound confidence level. Lifetime extrapolation is based upon the logarithmic time relation.

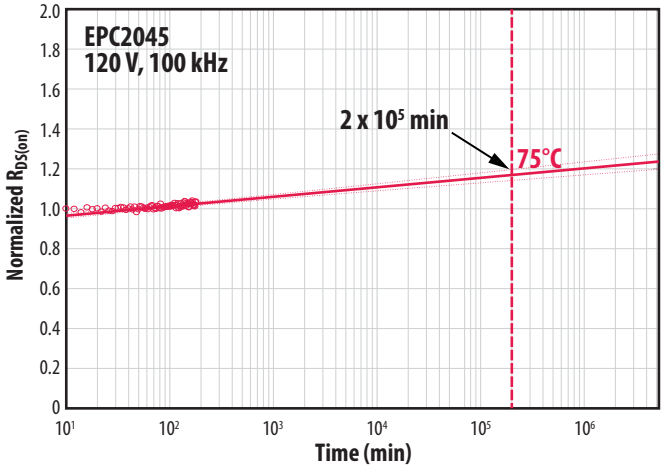


Figure 4-15: Evolution of $R_{DS(on)}$ of a representative EPC2045 device, a fifth-generation 100 V rated GaN transistor, tested at 120 V and 75°C. It is projected to exceed 20% $R_{DS(on)}$ shift at 2 x 10^5 minutes by considering 90% of upper bound confidence level.

It is noted that this is a more conservative estimated time of failure than the actual projected lifetime that is estimated to be nearly 1 x 10^6 minutes. By multiplying by 85%, it yields 1.7 x 10^5 minutes, representing the total lifetime when the DUT is off state biased continuously under 120 V and 75°C. When comparing with 25 years of expected overall lifetime, equivalent of 1.3 x 10^7 minutes, it translates to approximately 1.3% of total lifespan in mission. To add more margin, we rounded to 1% of 25 years. Now a total lifetime-based overvoltage specification of 1.3 x 10^5 minutes is developed.

To further validate this total time-based specification, the same testing conditions were applied to newer 100 V rated GaN products including EPC2218, EPC2071, EPC2302, and EPC2204. Figure 4-16 summarizes the testing results of the listed products, where they are all projected to outperform the 1.3 x 10^5 minutes of lifetime.

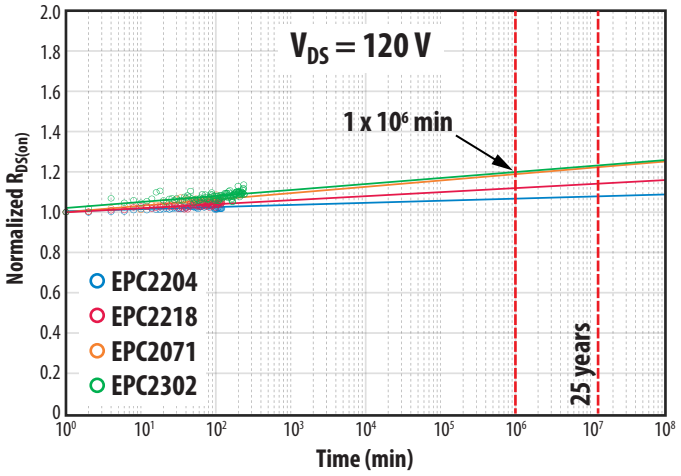


Figure 4-16: Evolution of $R_{DS(on)}$ of representative EPC2204, EPC2218, EPC2071, and EPC2302 GaN transistors, rated at 100 V and tested at 120 V and 75°C. They are projected to have less than 20% $R_{DS(on)}$ shift at a minimum of 1 x 10^6 minutes, significantly exceeding the 2 x 10^5 minutes lifetime based on EPC2045.

This total time-based specification can be scaled to a shorter duration that occurs repetitively within each switching cycle. Therefore, another way to specify this repetitive rating is to calculate the ratio of overvoltage duration of each cycle over the switching period, which is the 1% scaling factor that was initially discussed. This is equivalent to calculating the duty cycle of the overvoltage spike.

For instance, if a converter operates at 100 kHz, equivalent of 10 µs per switching period, it suggests that the GaN devices should withstand a repetitive 120 V overvoltage spike with a 100 ns duration in each switching cycle over 25 years of lifetime. This mathematical relation is demonstrated in Eq. 4-18 and further illustrated in Figure 4-17.

Eq. 4-18

$$Overshoot\ duty\ cycle = \frac{120\%\ Overvoltage\ Duration\ at\ 75°C\ (T_O)}{Switching\ Period\ (T_S)} \leq 1\%$$

where T_O is the overvoltage duration within each switching period and T_S is the switching period.

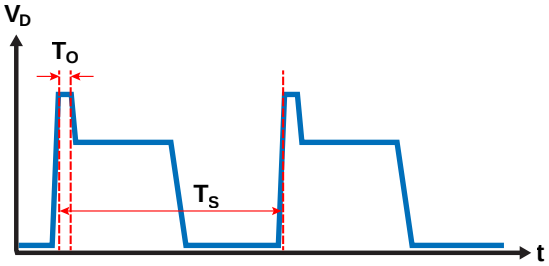


Figure 4-17: Illustration of the 1% overshoot duty cycle overvoltage specification. 1% is the ratio between T_O (overvoltage duration) and T_S (one switching period).

To verify this newly proposed overvoltage specification method, an unclamped inducive switching (UIS) circuit was developed [18]. Figure 4-18 shows the resulting overvoltage pulse that is generated by UIS.

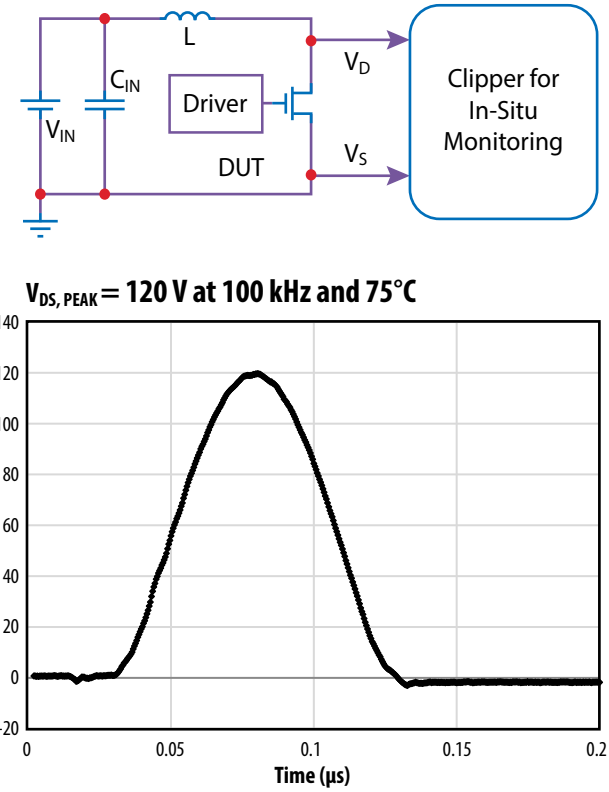


Figure 4-18: Simplified schematic of the unclamped inductive switching circuit and the resulting overvoltage pulse with $V_{DS,Peak}$ of 120 V under 100 kHz operating frequency.

A number of 100 V rated GaN transistors from different wafer lots are stressed by a 120 $V_{DS,Peak}$ overvoltage spike at 100 kHz operation frequency and 75°C junction temperature. Figure 4-19 shows that representative EPC2218 devices from three different wafer lots were tested to over billions of switching cycles showing very small dynamic $R_{DS(on)}$ shift [18].

The same physics-based lifetime model based on hot carrier trapping was applied to project the lifetime under such drain overvoltage stresses. The projection demonstrates the excellent robustness of GaN devices under 120% overvoltage stress over long-term continuous operation. At each switching cycle, the duration exceeding 100 $V_{DS,Max}$ is approximately 25 ns, lower than the 120 V peak overshoot voltage. At the end of 8 x 10^8 seconds (25 years), which equates to 8 x 10^{13} total pulses by multiplying with 100 kHz frequency, none of the DUTs surpassed the 20% $R_{DS(on)}$ shift failure criteria. Multiplying 25 ns by 8 x 10^{13} pulses gives 2 x 10^5 minutes, which is close to the estimated total lifetime of 1.3 x 10^5 minutes. The slight difference can be explained by the fact that the DUTs only reach the 120 V peak voltage for a very short portion of each pulse. The voltage waveform shown in Figure 4-18 is more representative of real time circuit applications.

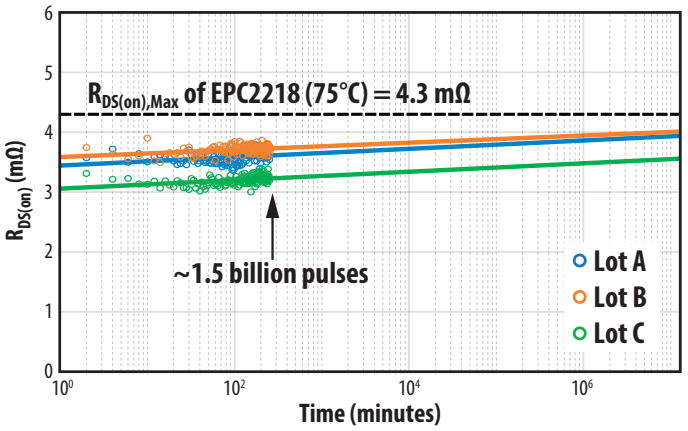


Figure 4-19: Evolution of dynamic $R_{DS(on)}$ of a representative EPC2218 DUTs from three different wafer lots under 120 $V_{DS,Peak}$ and 75°C UIS testing for more than 1.5 billion cycles.

Two additional representative 100 V-rated GaN transistors, EPC2204 and EPC2302 were tested under 120 $V_{DS,Peak}$ by UIS at 25°C shown in Figure 4-20. They were stressed for more than 6 and 10 billion pulses, respectively, where small dynamic $R_{DS(on)}$ drifts were measured. When projected to 25 years (1.3 x 10^7 minutes) and beyond, the dynamic $R_{DS(on)}$ of the DUTs are expected to be well below the maximum datasheet limit. The results further validated the proposed overvoltage specification.

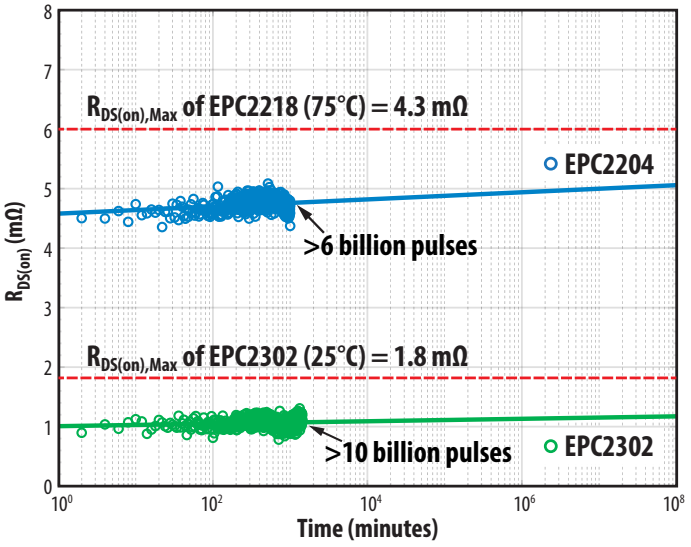


Figure 4-20: Evolution of $R_{DS(on)}$ shift of a representative EPC2204 and EPC2302 DUTs under 120 $V_{DS,Peak}$ UIS testing.

A repetitive drain overvoltage specification is proposed and validated by resistive load hard switching and unclamped inductive switching testing circuits. This duty cycle-based specification offers a more quantitative and easy-to-implement guideline for application engineers to design GaN devices. This work also demonstrates the extreme overvoltage robustness of GaN HEMTs.

4.3. Current Density Wear-out

4.3.1. Introduction to Current Density Wear-out Mechanisms

Thermal limits can become a concern for GaN devices when high current and high drain-source voltage occur simultaneously. Extensive robustness testing was conducted, and the results validated of safe operating area specified in the datasheet. For certain applications, the capability to withstand short circuit fault conditions is a must. Therefore, short circuit testing was performed, where GaN demonstrated excellent robustness under such extreme stress conditions. When devices are exposed to continuous high current at elevated temperatures, electromigration (EM) robustness becomes a common concern for customers. Thus, accelerated EM testing was conducted on power quad-flat no-leads (PQFN) devices that utilize copper pillars as the interconnects between the device and the package. Based on the EM testing results, a continuous current rating was developed for PQFN products, which also demonstrates excellent EM robustness. Lastly, a pulsed current rating specification was developed for GaN at various gate drive voltages and temperatures.

4.3.2. Safe Operating Area

Safe operating area (SOA) testing exposes the GaN transistor to simultaneous high current (I_D) and high voltage (V_{DS}) for a specified pulse duration. The primary purpose is to verify the transistor can be operated without failure at every point (I_D, V_{DS}) within the datasheet SOA graph. It is also used to probe the safety margins by testing to fail outside the safe zone. During SOA tests, the high-power dissipation within the die leads to a rapid rise in junction temperature and the formation of strong thermal gradients. For sufficiently high power or pulse duration, the device simply overheats and fails catastrophically. This is known as thermal overload failure.

In Si MOSFETs, another failure mechanism known as secondary breakdown (or Spirito effect [21]) has been observed in SOA testing. This failure mode, which occurs at high V_D and low I_D, is caused by unstable feedback between junction temperature and threshold V_{TH}. As the junction temperature rises during a pulse, V_{TH} drops, which can cause local current to rise. The rising current, in turn, causes temperature to rise faster, thereby completing a positive feedback loop that leads to thermal runaway and ultimate failure. The goal of this study is to determine if the Spirito effect exists in GaN transistors.

For DC, or long-duration pulses, the SOA capability of the transistor is highly dependent on the heatsinking of the device. This can present a huge technical challenge to assess the true SOA capability, often requiring specialty water-cooled heatsinks. However, for short pulses (< 1 ms), the heatsinking does not impact SOA performance. This is because on short timescales the heat generated in the junction does not have sufficient time to diffuse to any external heatsink. Instead, all the electrical power is converted to raising the temperature (thermal capacitance) of the GaN film and nearby

silicon substrate. As a result of these considerations, SOA tests were conducted at two pulse durations: 1 ms and 100 µs.

Figure 4-21 shows the SOA data of 200 V EPC2034C. In this plot, individual pulse tests are represented by points in (I_D, V_{DS}) space. These points are overlaid on the datasheet SOA graph. Data for both 100 µs and 1 ms pulses data are shown together. Green dots correspond to 100 µs pulses in which a part passed, whereas red dots indicate where a part failed. A broad area of the SOA was interrogated without any failures (all green dots), ranging from low V_{DS} all the way to $V_{DS,Max}$ (200 V). All failures (red dots) occurred outside the SOA, indicated by the green line in the datasheet graph. The same applies to 1 ms pulse data (purple and red triangles); all failures occurred outside of the datasheet SOA.

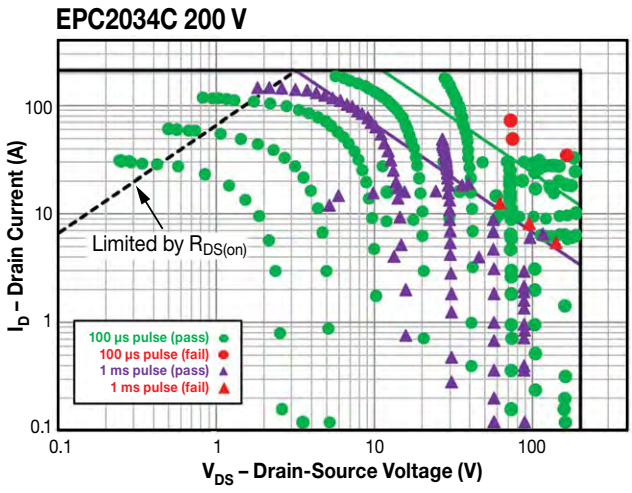
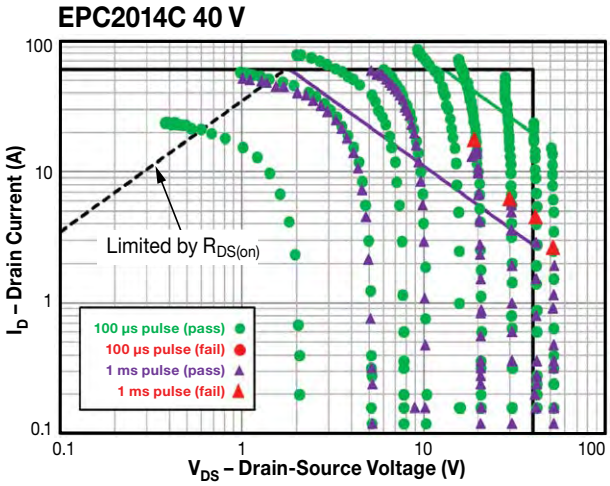
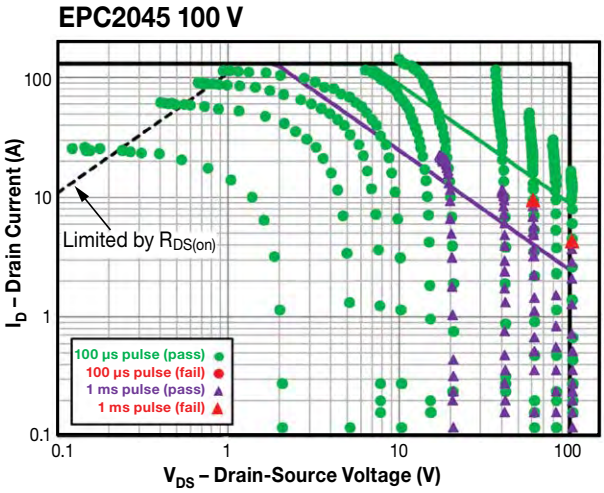


Figure 4-21: EPC2034C SOA plot. The "Limited by $R_{DS(on)}$" line is based on datasheet maximum specification for $R_{DS(on)}$ at 150°C. Measurements for 1 ms (purple triangles) and 100 µs (green dots) pulses are shown together. Failures are denoted by red triangles (1 ms) or red dot (100 µs). Note that all failures occur outside the datasheet SOA region.

Figure 4-22 provides SOA data for three more parts, EPC2212 (4th generation automotive 100 V), EPC2045 (5th generation 100 V), and EPC2014C (4th generation 40 V). In all cases, the datasheet safe operating area has been interrogated without failures, and all failures occur outside of SOA limits, often well outside the limits.

The datasheet SOA graph is generated with finite element analysis, using a thermal model of the device including all relevant layers along with their heat conductivity and heat capacity. Based on transient simulations, the SOA limits are determined by a simple criterion: for a given pulse duration, the power dissipation must be such that the junction temperature does not exceed 150°C before the end of the pulse. This criterion results in limits based on constant power, denoted by the 45° green (100 µs) and purple (1 ms) lines in the SOA graph. This approach leads to a datasheet graph that defines a conservative safe operating zone, as evidenced by the extensive test data in this study. In power MOSFETs, the same constant power approach leads to an overestimate of capability in the high voltage regime, where failure occurs prematurely due to thermal instability (Spirito effect). While the exact physics of failure is yet to be determined, the main outcome of this study is clear – GaN transistors will not fail when operated within their datasheet SOA.





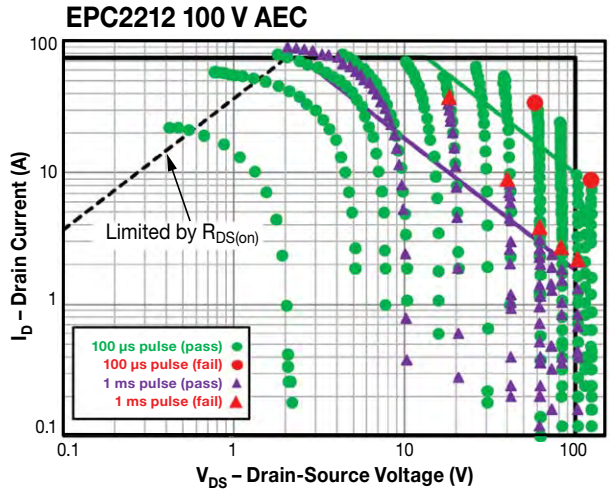


Figure 4-22: SOA results for EPC2014C, EPC2045, and EPC2212. Measurements for 1 ms (purple triangles) and 100 µs (green dots) pulses are shown together. Failures are denoted by red triangles.

4.3.3. Short-Circuit Robustness

Short circuit robustness refers to the ability of a FET to withstand unintentional fault conditions that may occur in an application while in the ON (conducting) state. In such an event, the device will experience the full bus voltage combined with a current that is limited only by the inherent saturation current of the transistor and the circuit parasitic resistance, which varies with the application and location of the fault. If the short-circuit state is not quenched by protection circuitry, the extreme power dissipation will ultimately lead to thermal failure of the transistor. The goal of short-circuit testing is to quantify the "withstand time" the part can survive under these conditions.

Typical protection circuits (e.g., de-saturation protection for IGBT gate drivers) can detect and react to over-current conditions in 2–3 µs. It is therefore desirable if the GaN transistor can withstand unclamped short-circuit conditions for about 5 µs or longer.

The two main test circuits used for short-circuit robustness evaluation are described in [22]. They are:

- Hard-switched fault (HSF): gate is switched ON (and OFF) with drain voltage applied.
- Fault under load (FUL): drain voltage is switched ON while gate is ON.

For this study, devices were tested in both fault modes and no significant differences in the withstand time were found. Therefore, the focus will be on FUL results for the remainder of this discussion. However, it is important to note that from HSF testing, GaN transistors did not exhibit any latching or loss of gate control that can occur in silicon based IGBTs [23]. This result was expected given the lack of parasitic bipolar structures with the GaN devices. Until the time the transistors fail catastrophically, the short circuit can be fully quenched by switching the gate LOW, an advantageous feature for protection circuitry design.

Two representative GaN transistors were tested:

1. EPC2203 (80 V): 4th generation automotive grade (AEC) device
2. EPC2051 (100 V): 5th generation device

These devices were chosen because they are the smallest in their product families. This simplified the testing owing to the high currents required for short-circuit evaluation. However, based on simple thermal scaling arguments, the withstand time is expected to be identical for other in-family devices. EPC2203 results cover EPC2202, EPC2206, EPC2201 and EPC2212; EPC2051 covers EPC2045 and EPC2053.

Figure 4-23 shows fault-under-load data on EPC2203 for a series of increasing drain voltages. With V_{GS} at 6 V (the datasheet maximum), and a 10 µs drain pulse, the device did not fail all the way up to V_{DS} of 60 V. Under these conditions, over 1.5 kW is dissipated in a 0.9 x 0.9 mm die. At the higher V_{DS}, the current is seen to decay over time during the pulse. This is a result of rising junction temperature within the device and does not signify any permanent degradation.

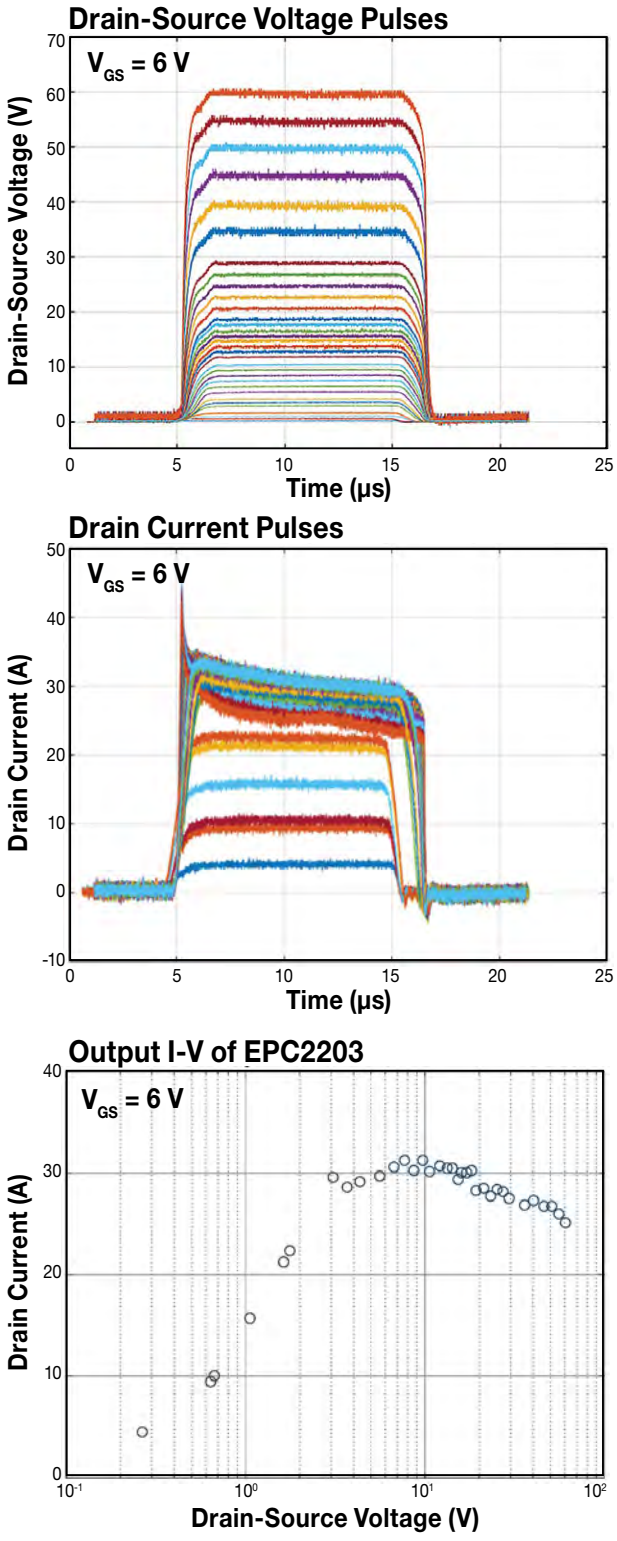


Figure 4-23: EPC2203 fault under load test (FUL) waveforms for a series of increasing drain voltages. Drain pulse is 10 µs and V_{GS} = 6 V. The device did not fail for this pulse width. In the V_{DS} vs. time plot (left), V_{DS} is Kelvin-sensed directly at the device terminals. In the I_{DS} vs. time plot (center), it is noted that I_{DS} decreases over time due to self-heating. Resulting output curve for this test sequence (right). Drain current is reported as the average current during the pulse. Drain current rolls over in the saturation region owing to device heating at higher V_{DS}.

Using a longer pulse duration (25 µs), the parts eventually fail from thermal overload. Representative waveforms are shown in Figure 4-24. The time of failure is marked by the abrupt sharp rise in drain current. After this event, the devices are permanently damaged. The withstand time is measured from the beginning of the pulse to the time of failure.

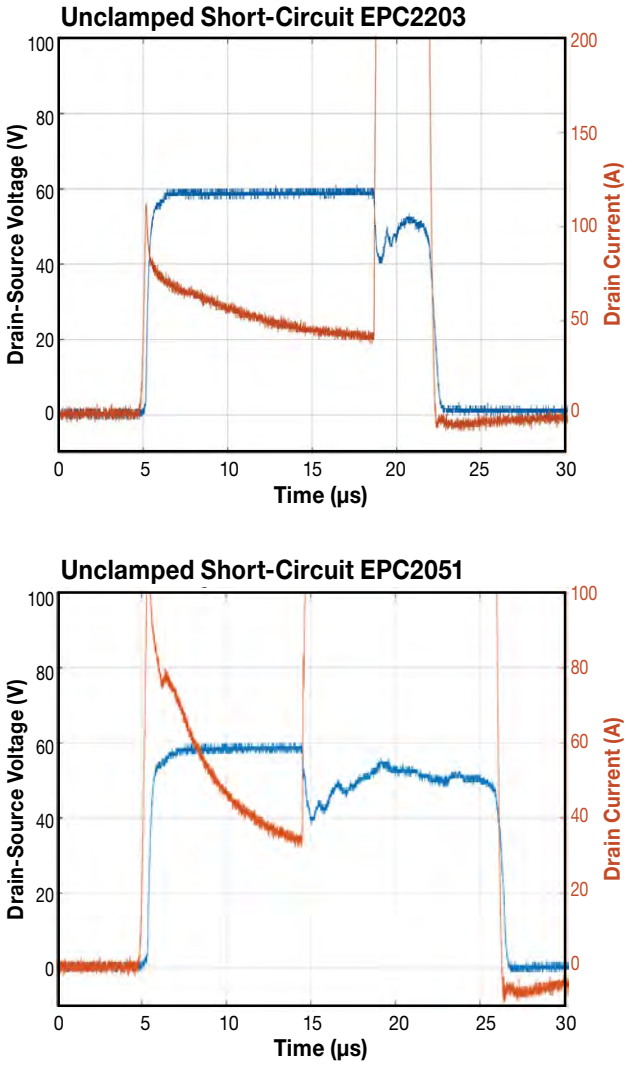


Figure 4-24: Fault-under-load test waveforms for a typical EPC2203 (left) and EPC2051 (right) at V_{DS} = 60 V, V_{GS} = 6 V, and a 25 µs drain pulse. The abrupt rise in drain current marks the time of catastrophic thermal failure.

To gather statistics on the withstand time, cohorts of eight parts were tested to failure using this approach. Table 4-1 summarizes the results. EPC2203 was tested at both 5 V (recommended gate drive) and 6 V ($V_{GS(max)}$), with mean withstand time of 20 µs and 13 µs, respectively. Note that the device survives less time at 6 V because of the higher saturation current. EPC2051 exhibited a slightly lower time-to-fail (9.3 µs) compared with the EPC2203 at 6 V. This is expected because of the more aggressive scaling and current density of 5th generation products. However, in all cases, the withstand time is comfortably long enough for most short-circuit protection circuits to respond and prevent device failure. Furthermore, the withstand time showed small part-to-part variability.

The lower rows in Table 4-1 provide pulse power and energy relative to die size. To gain insight into the relationship between these quantities and the time to failure, time-dependent heat transfer was simulated to determine the rise in junction temperature ΔT_J during the short-circuit pulse. The results are shown in Figure 4-25.

Short-circuit pulse V_{DS} = 60 V	**EPC2203 (Gen 4)**		**EPC2051 (Gen 5)**	
	V_{GS} = 6 V	V_{GS} = 5 V	V_{GS} = 6 V	V_{GS} = 5 V
Mean TTF (µs)	13.1	20.0	9.33	21.87
Std. dev. (µs)	0.78	0.37	0.21	2.95
Min. TTF (µs)	12.1	19.6	9.08	18.53
Avg pulse power (kW)	1.764	1.4	3.03	2.03
Energy (mJ)	23.83	27.6	27.71	42.49
Die area (mm²)	0.9025		1.105	
Avg power/area (kW/mm²)	1.95	1.55	2.74	1.84
Energy/area (mJ/mm²)	26.4	30.59	25.08	38.46

Table 4-1: Short-circuit withstand time statistics for EPC2203 and EPC2051

Note: Statistics derived from eight devices in each condition. Withstand times are tightly distributed around mean value. Average pulse power and energy correspond to a typical part within the population.

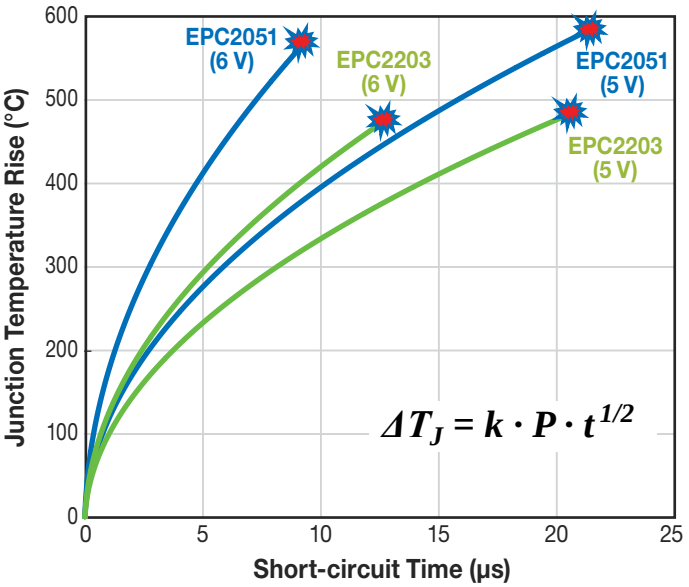


Figure 4-25: Simulated junction temperature rise versus time during the short-circuit pulses for both EPC2051 and EPC2203 at both 5 V and 6 V V_{GS}. Measured failure times are indicated by red markers. Note that EPC2203 fails catastrophically at a ΔT_J of around 475°C, whereas EPC2051 fails around 575°C. The simulated ΔT_J is well fit by a simple square root dependence on time (heat diffusion), as shown in the equation. P denotes the average power per unit area, and k = 6.73 x 10^{-5} $Km^2/Ws^{1/2}$.

The intense power density during the pulse leads to rapid heating in the GaN layer and nearby silicon substrate. Because the pulse is short and heat transfer is relatively slow, only a small thickness of semiconductor (< ~100 µm in depth) can help to absorb the energy. The temperature grows as the square root of time (characteristic of heat diffusion), and linearly with the pulse power. As can be seen in Figure 4-25, for EPC2203, both the 5 V and 6 V conditions fail at the same junction temperature rise of ~475°C. The same is true for EPC2051, where both conditions fail at the same ΔT_J of ~575°C. Three key conclusions stem from these results:

1. For a given device, the time to failure is inversely proportional to the power dissipation squared (P-2). This applies for short-circuit and SOA pulses of duration < ~1 ms.
2. The intrinsic failure mode resulting from high power pulses is directly linked to the junction temperature exceeding a certain critical value.
3. Wide bandgap eGaN devices can survive junction temperatures (>400°C) that are totally inaccessible to silicon devices owing to free-carrier thermal runaway.

4.3.4. Development of a Continuous Current Rating for PQFN GaN HEMTs

Copper pillars are used as the interconnects in the latest EPC's GaN HEMTs that utilize power quad flat no-leads (PQFN) packages. The copper pillar interconnects consist of two parts: a plated copper pillar and a solder cap that is mainly composed of Tin (Sn) with varying trace amounts of Silver (Ag), Gold (Au), and Copper (Cu) [24, 25, 26, 27, 28]. After the reflow process, the solder cap connects the die and the package and is typically considered as the limiting factor for the continuous current rating of GaN HEMTs. Electromigration (EM) has been identified as the primary wear-out mechanism, defined as the movement of atoms in a metal structure, leading to void formation [29,30]. Therefore, in this section, EM testing was conducted to determine the continuous current density limit for the copper pillars implemented in EPC's GaN HEMTs. Based on the test results, a continuous current rating is recommended with quantitative reliability implications.

The primary cause of EM is the electron "wind" generated from the transfer of momentum between conducting electrons and metal ions in the crystal. When the momentum surpasses the diffusion threshold that is governed by an activation energy [30,31] metal atoms can move and create voids. The Black's model is widely accepted to predict lifetime under EM wear-out mechanism, as shown in Eq. 4-19 [29,30].

$$MTTF = Aj^{-n}e^{\frac{Q}{kT}} \qquad \text{Eq. 4-19}$$

Where A is a constant, j is current density that is defined as current divided by the cross-sectional area of the copper pillar, n is an exponent, Q is the activation energy, k is the Boltzmann's constant at 8.62 x 10^{-5} eV/K, and T is the temperature in Kelvin unit.

The j^{-n} term in Black's equation models the solder wear-out, which is shown as void growth, is highly accelerated by current density. The initial formation of solder voiding, caused by EM degradation, reduces the cross-sectional area through which the current can flow, resulting in a further increase in current density. The increase in current density in turn further accelerates the solder void formation, which leads to a positive feedback loop. The $e^{Q/kT}$ term in Eq. 4-19 represents the thermal activation process of EM. Joule heating raises the junction temperature, which accelerates the movement of atoms resulting in more void formation. Both processes can lead to an open circuit due to void formation or electrical shorts caused by the melting of the metal interconnect. Since EM is a slow mechanism that can take years to develop under normal use conditions, testing under accelerated stress conditions is necessary to generate EM related failures within a reasonable timeframe.

The EM experiment consists of three parts, which include a device under test (DUT) card, a custom test chip, and a temperature chamber. The custom test chip was designed by following JEDEC standard, JEP154 [32]. The test setup is placed in a temperature chamber with the DUT card placed in the center. Two thermocouples were used: one mounted at the center of the oven to monitor the ambient temperature, and the other one is placed directly on the backside of the DUT, where the Si substrate is exposed. The test chip is covered with thermal putty and sandwiched between two copper heat sinks to maintain a constant temperature. The temperature difference between the copper pillar interconnect and the backside of the device, where the second thermocouple is placed, is calculated to be 0.64°C by using the $R_{th,JC}$ of 0.2°C/W and a total of 3.2 Watts of power dissipated at 125°C. The copper pillar interconnect of interest has an elliptical shape with an area of 5,271 um^2 and is soldered onto a copper lead frame that is molded into a PQFN package outline.

Test conditions of 27 kA/cm^2 at 125°C and 55 kA/cm2 at 150°C were selected, based on previous research studies focusing on copper pillar interconnects [26, 27, 28, 30]. A failure criterion of 10% resistance increase was adopted according to the recommendations in JEP154 [32]. Both test conditions yielded zero failures, which is consistent with various studies that focus on EM copper interconnects [24, 25, 26, 27, 28]. A current density power exponent of 2 has been frequently reported for copper pillar interconnects by various studies[24,27]. An activation energy of 1 eV is commonly accepted for SnAg solder cap through previous works [24, 25, 26, 27, 28]. By using the values of n=2 and Q=1 eV and assuming the time to failure of 870 hours with 0.1% failure rate, the constant A of the Black's equation is calculated to be 3.24. After determining the constant A, the lifetime at a 0.1% failure rate for any given temperature and current density can be calculated. The continuous current ratings of EPC's PQFN devices [24, 30] are based on a conservative EM current density limit of 13 kA/cm^2. By plugging in a current density of 13 kA/cm^2 and a junction temperature of 125°C into Eq. 4-12, 10 years of lifetime with 0.1% failure rate is projected.

4.3.5. Development of a Pulsed Current Rating for GaN HEMTs

In this section, a testing circuit was developed to systematically characterize the pulsed current rating of GaN HEMTs at various gate drive voltages and temperatures. After measuring a suite of GaN HEMTs, including various generation-5 and a representative generation-6 100 V drain-source rated devices, statistical analysis was performed to develop a recommendation for specifying the pulsed current rating of GaN HEMTs.

Figure 4-26 illustrates the schematics of the testing circuit. First, the low-side GaN device under test (DUT) is biased at the specified DC gate voltage (V_{GS}), while the high-side Si MOSFET (SIR500DP-T1-RE3) is turned on with a 25 µs gate pulse signal with an input voltage (V_{IN}) varying from 1 V to 5 V in steps of 0.5 V. During each pulse testing under a combination of V_{GS} and V_{IN}, measurements were taken using an oscilloscope to record the Kelvin-sensed drain-source voltage drop (V_{DS}) across the DUT, and the voltage drop across the 1 mΩ shunt resistor to calculate the drain-source current (I_{DS}).

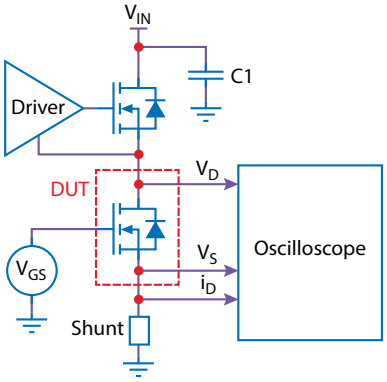


Figure 4-26: An illustration of the testing circuit to characterize the pulsed current rating under various V_{GS}, V_{DS}, and temperatures.

First, a matrix of generation-5 GaN HEMTs, including commercial-grade (EPC2051 and EPC2070) and automotive-grade (EPC2252 and EPC2204A), were tested under 5 V and 5.5 V gate drive voltages, as well as 25°C and 125°C device junction temperatures. The 125°C device junction temperature measurements were achieved by implementing a proportional–integral–derivative (PID) temperature controller directly mounted on the backside of the GaN devices, which have a low case-to-junction thermal resistance ($R_{\theta JC}$). Additionally, one generation-6 100 V drain-source voltage rated GaN HEMT (EPC2090) was tested under similar conditions.

Figure 4-27 shows a representative calculated current waveform under 5 V_{GS} and 3 V_{DS} at 25°C for EPC2252. The extracted pulsed current under such test conditions is obtained by averaging the measured current from 15 µs to 25 µs with a pulse width of 10 µs, as marked in Figure 4-27. Figure 4-28 summarizes all the averaged pulsed current measurements within a pulse width of 10 µs under various V_{DS}, V_{GS} and temperatures, where the vertical axis represents the measured current scaled by the corresponding gate width (Wg) for comparative analysis.

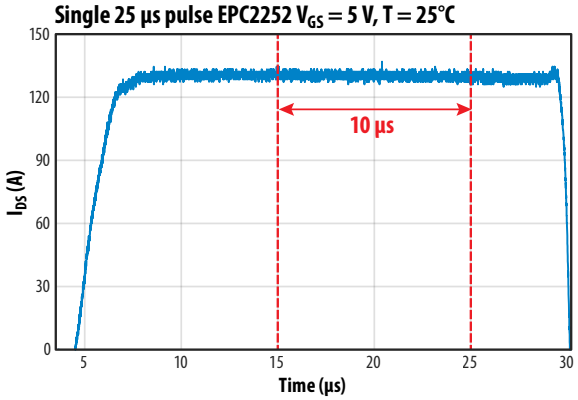


Figure 4-27: a representative drain-source current waveform of EPC2252 under 5 V_{GS} and 3 V_{DS} at 25°C, where a pulse width of 10 µs is used for pulsed current extraction.

As shown in Figure 4-28, a drain-source voltage of approximately 3 V is identified as the typical inflection point for 80 V or 100 V rated parts, at which the current conduction of GaN HEMTs transitions from linear region to saturation region. Therefore, 3 V_{DS} Kelvin-sensed measurement is used to quantify the pulsed current density at various V_{GS} and temperatures, with the results summarized in Table 4-2.

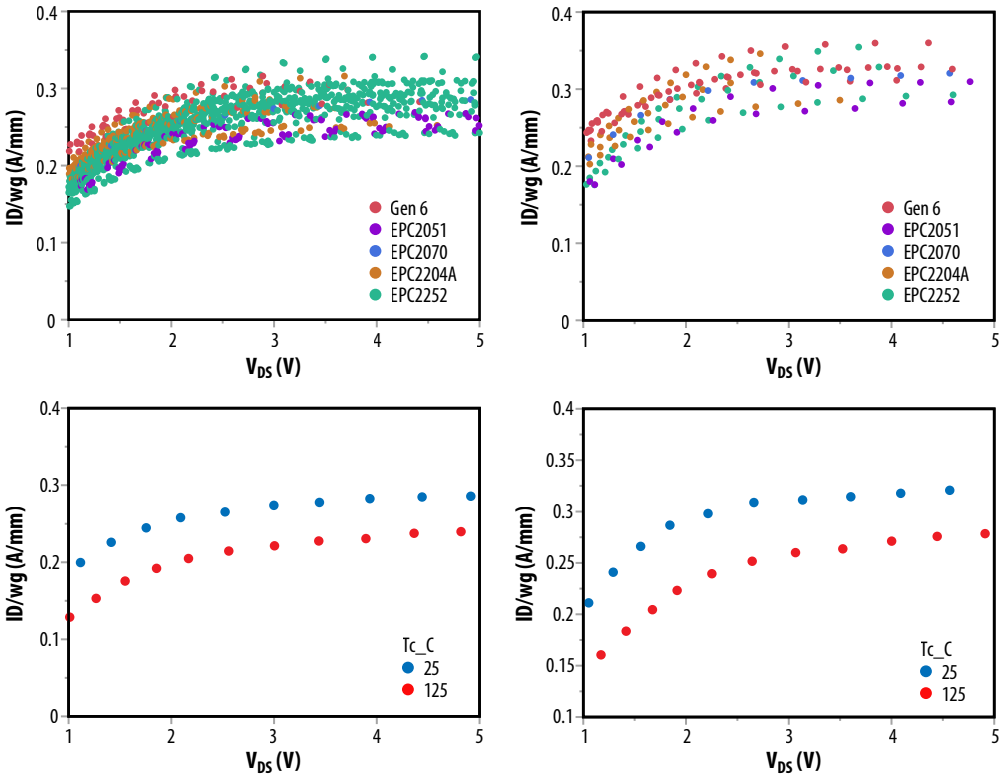


Figure 4-28: (a) pulsed current density scaled by gate width (Wg) vs. V_{DS} with a fixed V_{GS} of 5 V at 25°C; (b) pulsed current density scaled by gate width (Wg) vs. V_{DS} with a fixed V_{GS} of 5.5 V at 25°C; (c) comparison of the pulsed current density of a representative EPC2070 device under 25°C and 125°C with a V_{GS} of 5 V; (d) comparison of the pulsed current density of a representative EPC2070 device under 25°C and 125°C with a V_{GS} of 5.5 V.

Device Junction Temperature (°C)	I_D/W_g (A/mm) under V_{GS} = 5 V and V_{DS} = 3 V	I_D/W_g (A/mm) under V_{GS} = 5.5 V and V_{DS} = 3 V
25°C	0.27+/-0.02 (A/mm)	0.31+/-0.03 (A/mm)
125°C	0.22+/-0.03 (A/mm)	0.25+/-0.04 (A/mm)

Table 4-2 shows the mean pulsed current measurements with standard deviation, measured with a pulse width of 10 µs under various test conditions.

The three main conclusions of the pulsed current experiment are summarized as follows:

- At a 5 V gate drive and 25°C junction temperature, GaN HEMTs can consistently output a current density of more than 0.2 A/mm, even after considering three standard deviations.
- An approximately 15% increase in current output is expected when overdriving the gate from 5 V to 5.5 V.
- When the device junction temperature is increased from 25°C to 125°C, GaN HEMTs are expected to output approximately 20% less current due to the increased $R_{DS(on)}$.

In addition to the single pulsed current testing summarized in Table 4-2, long-term robustness testing was performed on four EPC2306 GaN HEMTs with a pulsed current density of ~0.33 A/mm at 25°C, which is equal to the mean current density of 0.27 A/mm plus three standard deviations. The resulting drain-source current is more than twice the maximum pulsed current rating of 197 A, specified in the datasheet. Figure 4-29 shows that after 100 million pulses with a testing frequency of 5 Hz, the $R_{DS(on)}$ of all DUTs remain well below the datasheet maximum specifications, suggesting the robustness of the GaN HEMTs under such extreme pulsed current stress conditions.

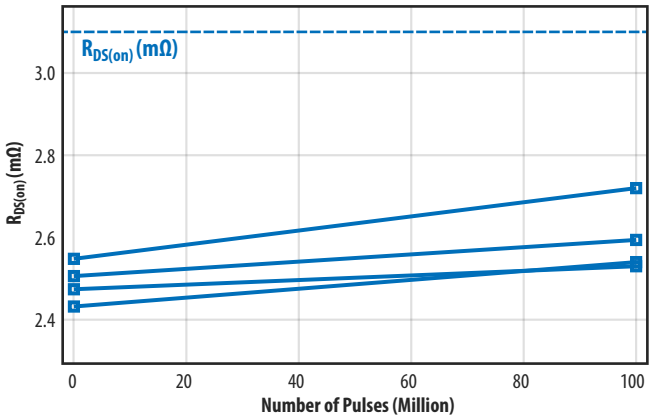


Figure 4-29: $R_{DS(on)}$ measurements before and after 100 million pulsed current stress with more than twice the maximum pulsed current rating of 197 A, specified in the datasheet.

4.4. Thermomechanical Wear-Out

4.4.1. Introduction to Thermomechanical Wear-Out Mechanisms

The primary wear-out mechanism responsible for thermomechanical stress is solder joint cracking, which occurs due to a mismatch in the coefficients of thermal expansion (CTE) between the DUT, the solder interconnects, and the PCB. Thermomechanical stress has emerged as a common concern in applications that experience frequent and large temperature swings. A comprehensive temperature cycling (TC) lifetime model is developed in this section, which includes device dimensions, bump size, TC test conditions, ramp rate, and various PCB properties. When the expected lifetime of chip scale packaged (CSP) devices is less than the customers' specifications, underfill with the right materials properties is recommended to improve TC lifetime. Lastly, the TC thermomechanical lifetime model is applied to power cycling (PC) stress for PQFN packaged devices, where the thermomechanical stress arises from the non-uniform temperature gradient between the PQFN devices and the PCB, resulting from the repetitive on-and-off operations of the devices.

4.4.2. Development of a Comprehensive TC Lifetime Model

In previous reliability reports [1, 15], the main wear-out mechanism mode under temperature cycling (TC) stress was identified as solder joint cracking [1]. Coefficient of thermal expansion (CTE) mismatch between the materials namely the device, solder and PCB is attributed as the fundamental cause of this wear-out mechanism. The CTE values of a typical FR4 PCB [33], a wafer level chip scale package (WLCSP) GaN-on-Si device [34], and SAC305 solder [35], are provided in Table 4-3. Figure 4-26 illustrates the resulting stress caused by CTE mismatch during temperature cycling testing. Figure 4-30 (a) shows the solder joint between the device and PCB in a neutral thermal stress position. As the temperature is lowered as in Figure 4-30 (b), the PCB with the higher CTE value contracts more than the GaN device, creating strain on the solder joints. Similarly, when the temperature increases in Figure 4-30 (c), the PCB undergoes more expansion than the device, again creating strain on the solder joints.

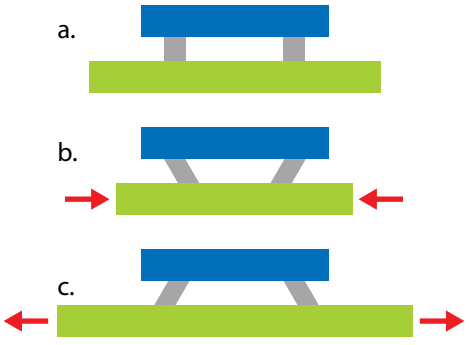


Figure 4-30 Illustration of stress on solder joints during temperature cycling.

Material	CTE (ppm/ °C)
Device	4
Solder	23
PCB (FR4)	18

Table 4-3: Common material coefficients of thermal expansion

In the following sections, a comprehensive TC lifetime model is developed by incorporating the effect of die size and bump shape, TC environmental test conditions and various PCB properties.

4.4.2.1. Modeling the Effect of Die Size and Bump Dimension

TC lifetime with respect to die size is typically modeled using the classic Coffin-Manson relation, where the devices under test (DUTs) are usually symmetrical in both the x and y directions [36]. Additionally, most of the solder joints presented in those studies are ball grid array (BGA), where all the bumps have an identical shape. Thus, distance-to-neutral point-based TC lifetime models are frequently adopted and have proven to be effective [37]. However, there is a lack of TC lifetime models that account for both asymmetrical die size and varying solder bump shapes with land grid array (LGA) solder bumps [38].

In this section, a suite of wafer level chip scale package (WLCSP) GaN devices with varying die size and bump shapes were evaluated for temperature cycling performance under a consistent assembly and TC testing condition. The Weibull distribution plots are shown in Figure 4-31, which includes EPC2206, EPC2071, EPC2069, EPC2218, EPC2204, EPC2152, and EPC2215. The temperature cycling experiment was constructed to ensure that the only variables are the device dimensions and bump shape. These devices were mounted on identical test PCB boards using identical solder (SAC305). The standoff height (i.e. the solder height after assembly) of ~130 µm was maintained during the assembly process. This was verified by performing physical cross-section of the assembled boards. The temperature cycle range was from -40°C to 125°C, with a ramp rate of 15°C/min and soak time of 10 minutes at the end points following industry standard JESD22-A104F [39]. After every temperature cycling interval, an electrical screening was performed to determine the number of failures, where exceeding datasheet limits was used as the failure criteria. A test-to-fail approach was adopted, where the devices are tested until a 50% failure rate is achieved. The failure distribution was analyzed using a two-parameter Weibull distribution for each device using maximum likelihood estimation (MLE) [40]. The resulting Weibull fits are indicated by solid lines in the graph of Figure 4-31, and the Weibull characteristics are in Table 4-4. The corner solder joint cracking was found to be the main wear-out mode throughout all devices analyzed by physical cross-sectioning and SEM inspection, establishing that wear-out of the smallest corner solder bump is the limiting factor for TC lifetime.

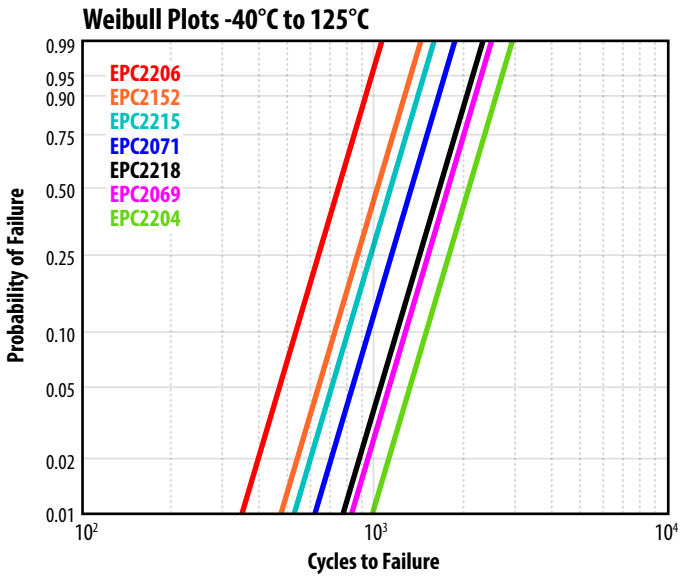


Figure 4-31: Weibull distribution fits to the experimental TC data of various CSP GaN products

Device	Weibull Shape Parameter	Characteristic Weibull Life (cycles)	Mean Time to Fail (cycles)
EPC2206	5.6	797	737
EPC2152	5.6	1085	1003
EPC2215	5.6	1199	1108
EPC2071	5.6	1416	1309
EPC2218	5.6	1764	1630
EPC2069	5.6	1880	1737
EPC2204	5.6	2389	2208

Table 4-4: Weibull statistics for tested devices

The Mean-Time-To-Fail (MTTF) data from the Weibull distribution, measured in number of cycles, were compared to die area to check for die size correlation with TC lifetime, as shown in Eq. 4-20.

$$MTTF = A(Die\ Area)^{-n} \qquad \text{Eq. 4-20}$$

where A is a constant, Die Area is the area of die by multiplying the length with the width and n is the exponent. The resultant fit is judged by a goodness-of-fit (R^2). A R^2 value of less than 0.7 indicates a poor fit, suggesting that die area alone is unable to provide a good correlation with TC lifetime by following the commonly accepted lifetime models in literatures [40, 41, 42].

The concept of "Maximum Distance from Neutral Point (DNP^{max})" is introduced as shown in Figure 4-32. During TC stress, the center point of the device experiences the least stress compared to extremities of the device. This center point is defined as the neutral point, the distance from the neutral point to the farthest extremity of solder bump is defined as DNP^{max}.

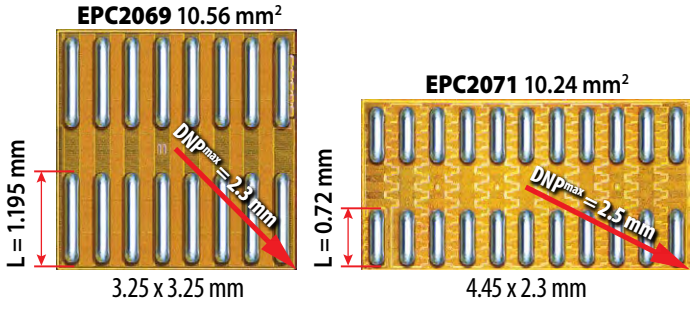


Figure 4-32: Example of gate length and DNP^{max} for EPC2069 and EPC2071.

By combining Norris-Landzberg modified Coffin-Manson TC lifetime model [44] and the concept of DNP^{max}, the MTTF can be modeled by Eq. 4-21, as reported by multiple researchers [45].

$$MTTF = A(DNP^{max})^{-n} \qquad \text{Eq. 4-21}$$

The best fit to Eq. 4-21 yielded an R^2 value of 0.79, slightly improved compared with simply using the device area. However, it is still not considered a very good fit.

Failure analysis established the gate solder joint cracking at the device corner as the limiting factor for TC performance. A longer gate bump likely indicates a longer time to failure under TC stress and vice versa. Figure 4-31 and 4-32 show that different device sizes also have varying length of the gate solder bump. Therefore, the corner gate bump shape should also be considered along with the DNP^{max} for a more accurate TC lifetime model development. Because the gate bump width is similar for all devices studied, the bump length, denoted as L, is the primary parameter that is included in the following discussions. Thus, the length of solder bump L is factored into DNP^{max}, and effective DNP (DNP^{eff}) is defined in Eq. 4-22.

$$DNP^{eff} = DNP^{max} + a \cdot L \quad \text{Eq. 4-22}$$

The resulting fit is shown in Figure 4-33 and results in an R^2 value of 0.99 using gate length factor a = -0.65, and power exponent n = 1.4.

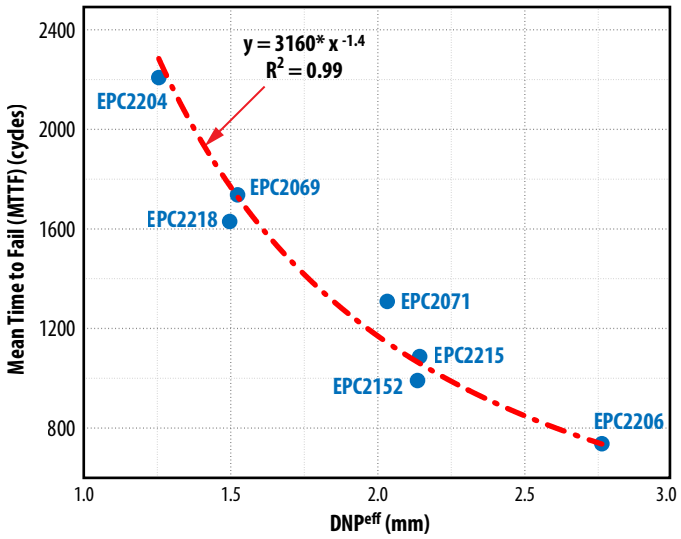


Figure 4-33: Measured MTTF under TC conditions of -40°C to 125°C vs. the effective DNP (DNP^{eff}) of 7 different devices with varying die dimensions and bump shape, where the red dash line, based on Eq. 4-16, provides an excellent fit to the measured MTTF.

The fitted power exponent of 1.4 shown in Figure 4-33 is consistent with other literature results [46, 47], where exponents between 1 and 2 are frequently reported in SAC305 solder joint cracking failures under TC stress with similar test conditions.

In summary, a TC lifetime model is proposed considering the device size and corner gate bump shape,

$$MTTF = A(DNP^{eff} - 0.65 \cdot L)^{-n} \quad \text{Eq. 4-23}$$

This study establishes a temperature cycling lifetime model based on solder joint cracking caused by CTE mismatch from materials which takes into consideration the varying dimensions of both die and solder joints.

COMSOL finite element analysis (FEA) simulations were carried out to validate the TC lifetime model presented in Eq. 4-23. Anand viscoplasticity model for the SAC305 solder was implemented in COMSOL, simulating solder's plasticity and creep behavior during temperature cycling [47, 48]. Hence, the energy dissipation density of the solder bumps can be calculated based on the area of stress-strain hysteresis loops, denoted as ΔW. Deveraux's energy-based fatigue model was subsequently used to calculate the MTTF, quantifying when the solder joint cracking initiates and eventually propagates through the entirety of the gate bump length, L, shown in Eq. 4-24 [48, 49].

$$MTTF = K_1 \Delta W^{K_2} + \frac{L}{K_3 / \Delta W^{K_4}} \quad \text{Eq.4-24}$$

where the first term, $K_1\Delta W^{K_2}$ represents the crack initiation lifetime and the second term, $\frac{L}{K_3/\Delta W^{K_4}}$, models the crack growth lifetime. K_1, K_2, K_3, and K_4 are fitting coefficients.

Table 4-5 shows that the simulated MTTF is within +/-10% error margin compared to experimental MTTF. This further validates the effectiveness of the proposed TC lifetime model in Eq. 4-17 that includes device dimensions and the critical corner gate bump length (L).

Product Area	Experimental MTTF (cycles)	COMSOL Simulated MTTF (cycles)	Δ_{MTTF}
EPC2619	2208	2284	+3%
EPC2218	1630	1537	-6%
EPC2069	1737	1849	+6%
EPC2206	737	792	+7%

Table 4-5: A summary of the modeled MTTF using COMSOL FEA vs. the experimental measured MTTF column, showing that the simulation agrees with the experimental data within 10%.

4.4.2.2. Modeling the Effect of TC Test Conditions

In this section, a comprehensive TC lifetime equation is developed to model various TC test conditions, including the temperature difference between the hot and cold temperature extremes (ΔT), the hot temperature extreme (T_{Max}), the ramp rate (R) and the dwell time at temperature extremes (t_{Dwell}).

First, TC experiments with different ΔT were performed on EPC2218A WLCSP devices, with both test legs using similar ramp rate (R) and dwell times (t_{Dwell}) at the temperature extremes. After every temperature cycling interval, electrical screening was performed, in which exceeding datasheet limits was used as the failure criteria. The two test conditions are TC1: –40°C to 125°C with ΔT = 165°C and T_{Max} = 125°C, and TC2: –40°C to 105°C with ΔT = 145°C and T_{Max} = 105°C. Figure 4-34 shows the Weibull distribution analysis of the two TC experiments, where TC1 with a larger ΔT and T_{Max} accelerated TC failures more than TC2. Therefore, the Norris-Landzberg lifetime model was used and shown in Eq. 4-25 [44].

$$N_{TC} = A \cdot f^{\alpha} \cdot \Delta T^{-\beta} \cdot exp\left(\frac{E_a}{kT_{Max}}\right) \quad \text{Eq.4-25}$$

where N_{TC} is the number of TC cycles to fail, f is the cycling frequency, describing the total number of cycles per day, and α is the cycling frequency exponent, which is typically specified as 1/3 [50, 51, 52, 53, 54]. ΔT defines the difference between T_{Max} and T_{Min} within one cycle and β is the temperature range exponent, typically dependent upon the solder type and properties. Since SAC305 solder is used in this study, a value of ~2 for β is used based on literature [50, 51, 52, 55, 56]. The last variable is the exponential term in Eq. 4-25, which is an Arrhenius term focusing on the creep mechanism at the maximum temperature, T_{Max}. E_a is the activation energy, k is the Boltzmann constant, 8.62 x 10^{-5} eV/K. The activation energy (E_a) at T_{Max} was calculated to be ~0.2 eV, based on Table 4-6.

This study forms the basis for the temperature-cycling reliability analysis of solar and DC-DC converters presented in Sections 5.1.6 and 5.2.5, respectively.

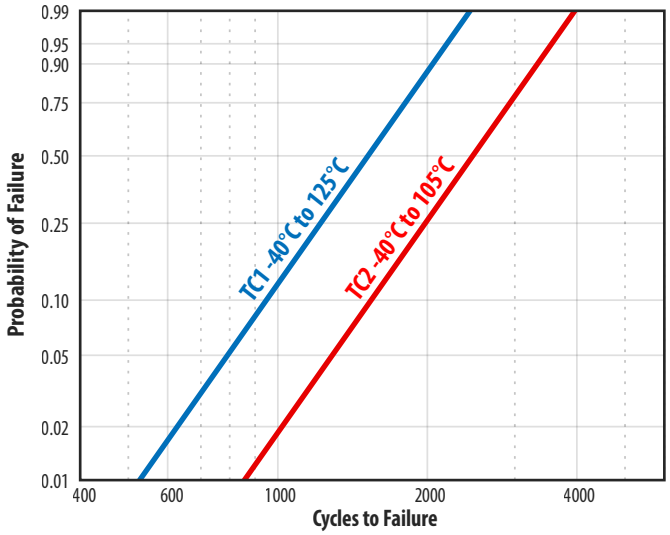


Figure 4-34. Weibull plots of temperature cycling results for EPC2218A under TC1 and TC2 test conditions, where devices are mounted on 2-copper-layer PCBs.

TC Condition	T_{Min} (°C)	T_{Max} (°C)	Characteristic Weibull Life	MTTF (cycles)
TC1 without underfill	165	40	36	1505
TC2 without underfill	145	30	48	2430
TC1 with underfill	165	40	36	7230 (Lower bound confidence level)

Table 4-6: Temperature cycling profile and MTTF determined by Weibull plots

In the Norris-Landzberg model, the frequency term ($f^{-\alpha}$) combines both the ramp rate and dwell time into a single term with the same exponent, which assumes that these two components have the same behavior and weight in relation to the MTTF. However, in many cases, the experimental results contradict the model's projections [57]. Therefore, a further set of TC experiments was conducted to deconvolute the frequency term in Eq. 4-25 into separate ramp rate term (R) and dwell time term (t_{Dwell}), each with its own power exponent. EPC2206 was used as the DUTs.

In this new study, the ramp rate (R) was varied from an average of 4°C/min to 14°C/min using a single-zone environmental TC chamber, while all other TC testing parameters remained consistent. Figure 4-35 shows the Weibull plots of the two TC experiments with different ramp rates.

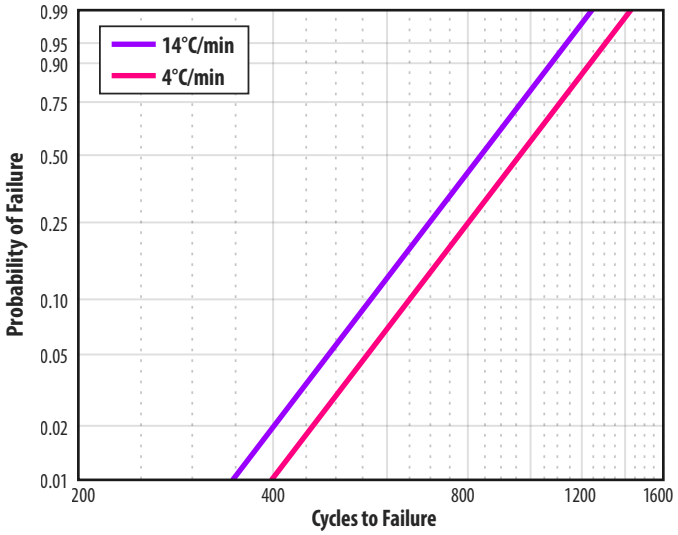


Figure 4-35. Weibull plots of EPC2206 with two different ramp rates in the temperature profile. The slow ramp rate = 4°C/min and the fast ramp rate = 14°C/min.

Eq. 4-26 is proposed to further define the ramp rate (R) and dwell time (t_{Dwell}), based on Eq. 4-25.

$$N_{TC} = A \cdot R^{a} \cdot {t_{Dwell}}^{b} \cdot \Delta T^{-\beta} \cdot \exp\left(\frac{Ea}{kT_{Max}}\right) \quad \text{Eq.4-26}$$

Figure 4-35 shows that the MTTF of the fast TC chamber (R = 14°C/min) is 829 cycles, which ~13% less than that of the slow TC chamber (R = 4°C/min), with a MTTF of 952 cycles. Therefore, the ramp rate exponent, a, is estimated to be -0.134. The dwell time exponent, b, is -1/3 based on literature [50, 51, 52, 53, 54], suggesting that longer dwell time at TC temperature extremes lead to lower TC lifetime. Therefore, Eq. 4-26 can be simplified to Eq. 4-27 in terms of the TC ramp rate (R).

$$N_{TC} \propto R^{-0.134} \quad \text{Eq.4-27}$$

Figure 4-36 shows the normalized TC lifetime as a function of the TC ramp rate under -40°C to 125°C test conditions, with all TC lifetimes normalized to that of the 15°C/min ramp rate. 15°C/min is used because it is the most commonly referenced TC ramp rate in the JEDEC standard [39] for evaluating the reliability of solder interconnects. Therefore, users can refer to Figure 4-36 to extrapolate the TC lifetime at different TC ramp rates based on the existing TC data.

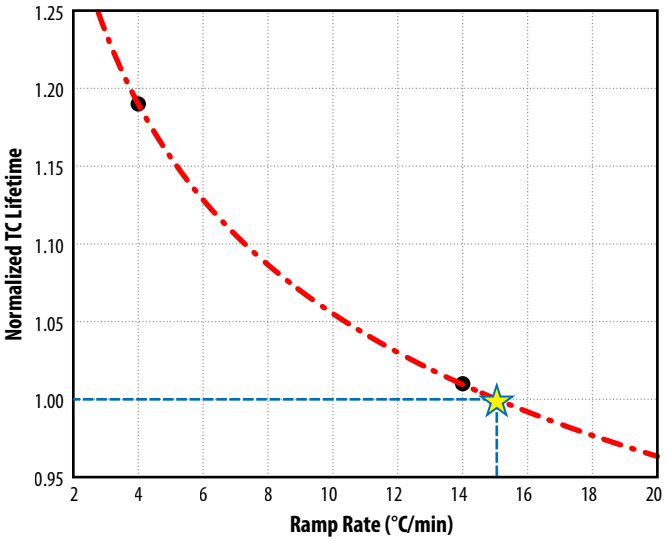


Figure 4-36: Normalized TC lifetime vs. ramp rate on a 2-layer PCB under TC condition of -40°C to 125°C, with all TC lifetimes normalized to 15°C/min ramp rate.

Finite Element Analysis (FEA) simulations were performed using COMSOL Multiphysics® to investigate the underlying mechanism responsible for TC ramp rate effect on solder joint lifetime. The stress-strain hysteresis loops for both ramp rate groups are illustrated in Figure 4-37. The higher ramp rate group exhibits higher stress levels compared to the slow group, resulting in increased energy dissipation density and, consequently, a shorter TC lifetime. The higher strain rate under higher ramp rate can lead to a more significant strain hardening effect in the SAC305 solder, thereby generating elevated stress levels within the solder [48]. Therefore, the simulation results predict that the MTTF at R = 14°C/min is 10.1% lower than that at R = 4°C/min, which agrees reasonably well with the experimental difference of 13%.

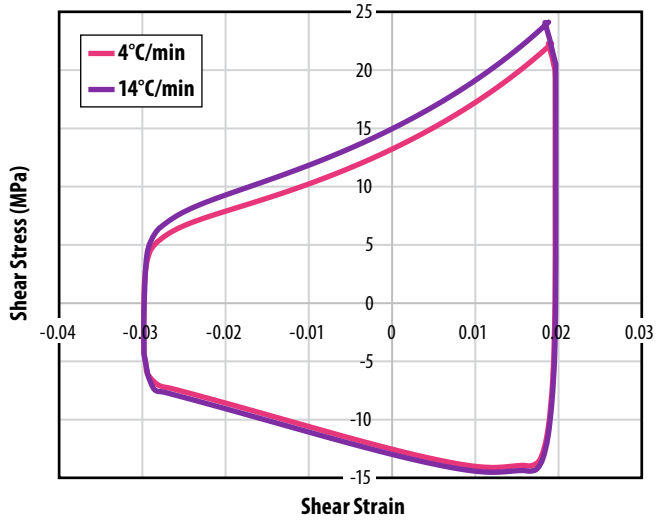


Figure 4-37: Stress-strain hysteresis loop of the slow and fast ramp rate groups.

4.4.2.3. Modeling the Effect of PCB Properties

High-density power modules often utilize high-layer count and thick printed circuit boards (PCBs). Such implementations raise concerns about solder joint reliability during TC due to the increased stiffness of these complex PCBs. The influence of PCB properties on solder joint lifetime under TC stress can be modeled by Clech's "board thickness" model [41]. Clech's model is developed by modeling the mechanical coupling between component and PCB from first principles. Although it is commonly referred to as the "board thickness" model, it is a comprehensive model that accounts for all critical parameters involving the component, board, and assembly. Based on Clech's model, the overall lifetime, N_{Total}, consists of three parts of life which associate three different mechanical coupling mechanisms.

The first part, N_1, is the lifetime that is characterized by the in-plane tensile shear force, acting on the device. Figure 4-38 illustrates the evolution of the dimensional changes of a device and a PCB when the ambient temperature increases from a low temperature, where the stress on the solder joints is neutral, to the hot temperature extreme where the device expands significantly less than the PCB due to the CTE mismatch. As a result, the solder joints are stretched laterally as shown in Figure 4-38. N_1 represents the in-plane tensile stiffness of the mounted device as shown by the green arrow in Figure 4-38. Eq. 4-28 specifies the lifetime caused by such in-plane stencil shear force.

$$N_1 = \frac{F}{\Delta\alpha^2} \times \frac{1-\gamma_{QFN}}{E_{QFN}h_{QFN}} = \frac{F}{\Delta\alpha^2} \times C_1 \qquad \text{Eq. 4-28}$$

where F is a constant for a specific device-PCB system and under a given TC stress condition, $\Delta\alpha$ is the CTE mismatch between the device and PCB, γ_{Dev} is the Poisson's ratio of the device, E_{Dev} is its Young's modulus, and h_{Dev} is the height of the device. C_1 is denoted as the axial compliance of the device, $\frac{1-\gamma_{QFN}}{E_{QFN}h_{QFN}}$.

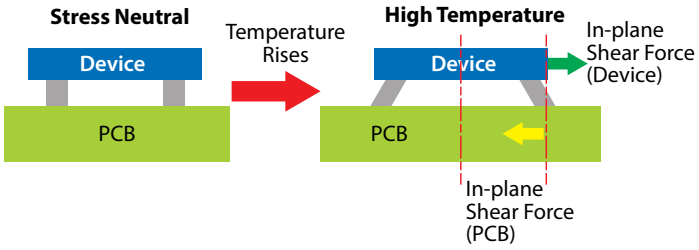


Figure 4-38: Illustration of the in-plane tensile shear forces acting on the device and PCB.

The second term, N_2, is controlled by the in-plane tensile shear force that acts on the PCB as highlighted by the yellow arrow in Figure 4-38. Eq. 4-29 characterizes the corresponding lifetime that is related to such tensile stiffness of the PCB.

$$N_2 = \frac{F}{\Delta\alpha^2} \times \frac{1-\gamma_{PCB}^2}{2E_{PCB}h_{PCB}} = \frac{F}{\Delta\alpha^2} \times C_2 \qquad \text{Eq. 4-29}$$

where F and $\Delta\alpha$ are the same as in Eq. 4-28, γ_{PCB} is the Poisson's ratio of the PCB, E_{PCB} is its Young's modulus, and h_{PCB} is the PCB thickness. C_2 is defined as the axial compliance of the PCB, $C_2 = \frac{1-\gamma_{PCB}^2}{2E_{PCB}h_{PCB}}$.

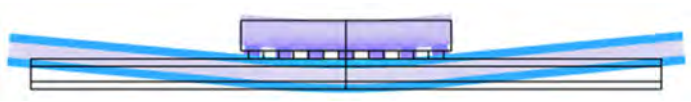


Figure 4-39 COMSOL FEA simulation results illustrate the flexural bending between the device and PCB.

Lastly, N_3 represents the bending moments of the bimetallic strip of the device and PCB, as shown in Eq. 4-30. Figure 4-39 shows the FEA simulation result of such bending motion. This part of lifetime, N_3, is dominated by the flexural modulus of the device and the PCB.

$$N_3 = \frac{F}{\Delta\alpha^2} \times \frac{H^2}{\frac{E^f_{QFN} h^3_{QFN}}{12(1-\gamma_{QFN})} + \frac{E^f_{PCB} h^3_{PCB}}{6(1-\gamma^2_{PCB})}} = \frac{F}{\Delta\alpha^2} \times C_3 \qquad \text{Eq. 4-30}$$

Where E^f_{Dev} and E^f_{PCB} are the flexural Young's modulus of the device, respectively. C_3 is the bending compliance of the bimetallic strip assembly of the device and PCB, $\frac{H^2}{\frac{E^f_{QFN} h^3_{QFN}}{12(1-\gamma_{QFN})} + \frac{E^f_{PCB} h^3_{PCB}}{6(1-\gamma^2_{PCB})}}$

and H is further defined by Eq. 4-31.

$$H = \frac{h_{QFN}}{2} + h_{standoff} + \frac{h_{PCB}}{2} \qquad \text{Eq. 4-31}$$

where $h_{Standoff}$ is the standoff height of the solder joint post-assembly.

Therefore, the total lifetime N_{Total} is determined by the sum of all three parts, as shown in Eq. 4-32.

$$N_{Total} = N_1 + N_2 + N_3 = \frac{F}{\Delta\alpha^2} \times (C_1 + C_2 + C_3) \qquad \text{Eq. 4-32}$$

Previous reliability reports showed that N_3, representing the bending motion interacting between the device and the PCB, dominates the total lifetime, N_{Total} [47]. Since the h_{PCB} used in high-power density applications is significantly thicker than both h_{Dev} and $h_{Standoff}$, H is essentially equal to h_{PCB}. Therefore, N_{Total} can be simplified to Eq. 4-33.

$$N_{Total} \propto \frac{A}{B+C \times h_{PCB}} \qquad \text{Eq. 4-33}$$

where A, B and C are constants that depend on the material properties of the PCB, the device and the solder joints post-assembly. Eq. 4-33 suggests that the TC lifetime is inversely proportional to the PCB thickness, assuming all other parameters remain constant as the PCB thickness decreases.

Similar accelerated TC experiment was conducted on EPC2218 (identical to EPC2218A in package) mounted on a 16-copper-layer PCB with a total thickness of 3.2 mm. The TC test conditions and assembly of the 16-layer PCB were consistent with those of the 2-copper-layer PCBs, with TC1 condition: -40°C to 125°C. Weibull distribution analysis found that the MTTF of the 16-layer in TC experiment decreased by approximately 40% compared to the MTTF of the 2-layer PCB, which is consist with the projection from the model in Eq. 4-33. Figure 4-40 shows the TC lifetime extrapolation from Eq. 4-33 as a function of the number of the copper layers within the PCB, based on three assumptions. First, the PCB thickness scales linearly with the number of copper layer with the copper thickness per layer being two Oz, approximately 70 µm. Second, the prepreg material is made of standard FR4 with a CTE of 18 ppm/°C. Lastly, the modulus and CTE mismatch between device and PCB remain constant as the number of copper layers decreases, indicating the total Cu/FR4 ratio stays consistent.

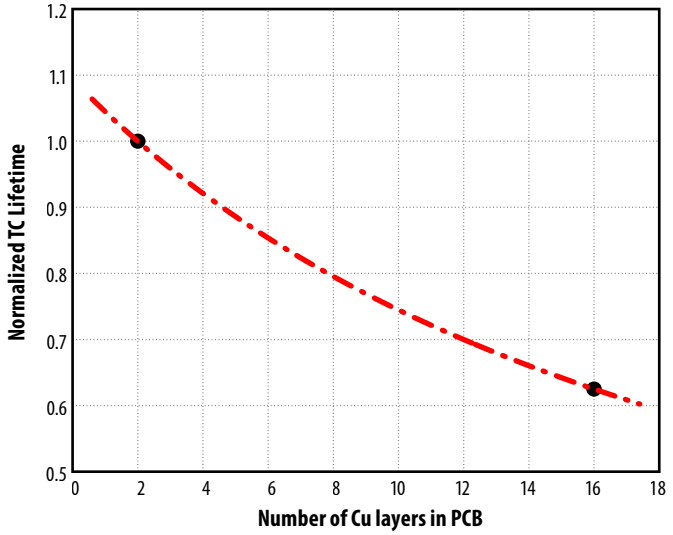


Figure 4-40: TC Lifetime vs. number of copper layer in PCB, where all MTTF is normalized to MTTF of 2 Cu-layers PCB.

4.4.3. Criteria for Choosing a Suitable Underfill

The selection of underfill material should consider a few key properties of the material as well as the die and solder interconnections. First, the glass transition temperature of the underfill material should be higher than the maximum operating temperature in application. Also, the CTE of the underfill needs to be as close as possible to that of the solder since both will need to expand/contract at the same rate to avoid additional tensile/compressive stress in the solder joints. As a reference, typical lead-free SAC305 and Sn63/Pb37 have CTEs of approximately 23 ppm/°C. Note that when operating above the glass transition temperature (Tg), the CTE increases drastically. Besides Tg, and CTE, the Young (or Storage) Modulus is also important. A very stiff underfill can help reduce the shear stress in the solder bump, but it increases the stress at the corner of the device, as it will be shown later in this section. Low viscosity (to improve underfill flow under the die) and high thermal conductivity are also desirable properties.

Manufacturer	Part number	CTE (ppm/ °C)			Storage modulus (DMA) at 25°C (N/mm²)	Viscosity at 25°C	Poisson's Ratio	Volume Resistivity	Thermal Conductivity	Dielectric Strength
		Tg (TMA) [C]	Below Tg	Above Tg						
HENKELS LOCTITE	ECCOBOND-UF 1173	160	26	103	6000	7.5 Pa*S				
NAMICS	U8437-2	137	32	100	8500	40 Pa*S	0.33	>1E15 Ω-cm	0.67 W/m · K	
NAMCIS	XS8410-406	138	19	70	13000	30 Pa*S				

Table 4-7: Recommended underfill material properties for WLCSP GaN devices

The main guidelines for choosing an underfill for use with GaN transistors are listed below:

- Underfill CTE should be in the range of 16 to 32 ppm/°C, centered around the CTE of the solder joint (24 ppm/°C). Lower values within this range are preferred because they provide better matching to the die and PCB.
- Glass transition temperature (Tg) should be comfortably above the maximum operating temperature. When operated above Tg, the underfill loses its stiffness and ceases to protect the solder joint.
- Young's (or Storage) modulus in the range of 6–13 GPa. If the modulus is too low, the underfill is compliant and does not relieve stress from the solder joints. If it is too high, the high stresses begin to concentrate at the die edges.

To better understand the key factors influencing thermo-mechanical reliability when using underfills, finite element simulations of EPC2206 under temperature cycling stress were conducted. Figure 4-41 shows the simulation deck used for this analysis. The die is placed on a 1.6 mm FR4 PCB, and the temperature change is ΔT = +100°C above the neutral (stress free) state. Two key underfill parameters were varied: Young's modulus and CTE. As shown in the figure, stress is analyzed along the cut line shown, providing visibility into the stress within the solder bars, die, and underfill.

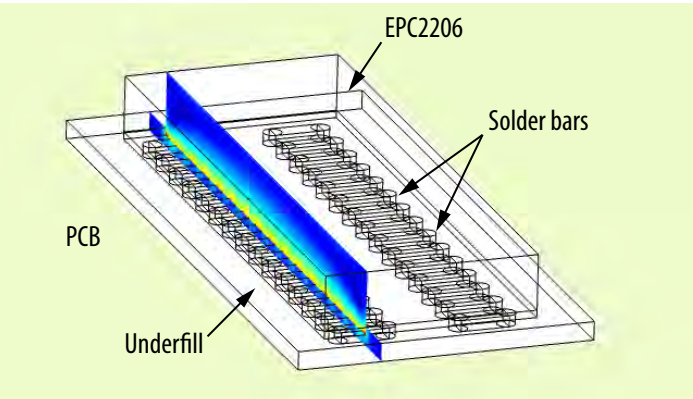


Figure 4-41: Simulation deck for finite element analysis of stresses inside EPC2206 under temp cycling stress. Die with underfill sitting on 1.6 mm FR4 PCB. Stress is analyzed along cut line shown.

Figure 4-42 shows the von Mises [58] peak shear stress in the edge-most solder bar along the cutline. For clarity, only stress in the solder bar is shown. In addition, mechanical deformations are exaggerated by 20 times in order to illustrate the shear displacement in the joint. Four distinct underfill conditions are simulated by changing the Young's modulus (E) or the CTE of the underfill. As can be seen, the solder bar in the no underfill case has by far the most extreme shear stress and deformation. The addition of underfill significantly alleviates stress from the joint. Higher Young's modulus reduces this stress further. For underfills with poor CTE matching to the solder joint, stress can also build up in the joint.

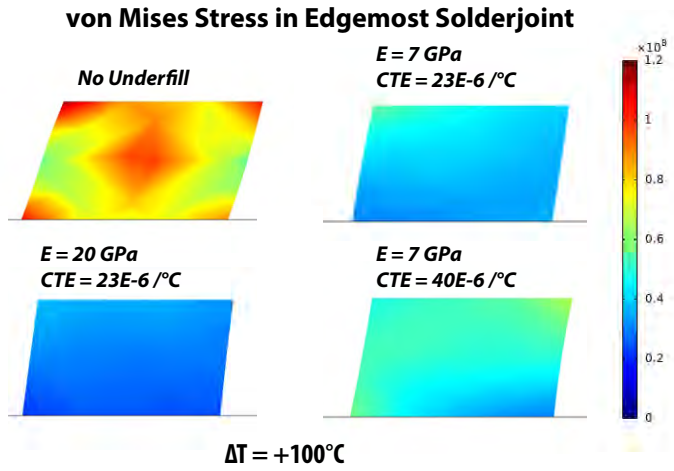


Figure 4-42: Von Mises (peak shear stress) in the edge-most solder bar under a temperature cycle change of ΔT = +100°C. Four different underfill conditions are simulated, with changing Youngs modulus (E) of the underfill, and different CTE as well. Note that mechanical deformation has been exaggerated by 20x in all cases.

Figure 4-43 shows the same four conditions, but this time the von Mises stress is shown in both the die and underfill. The high Young's modulus cases show low stress in the solder joint, but high stress inside the die and underfill near the die edge. These high stresses can lead to cracking and ultimate failure inside the device. FEA analysis shows that there is an optimal Young's modulus in the range of ~6 to 13 GPa, providing a good compromise between protecting the solder joint and protecting the die edge. Regarding CTE, the analysis shows that high underfill CTE (> 32 ppm/°C) should be avoided.

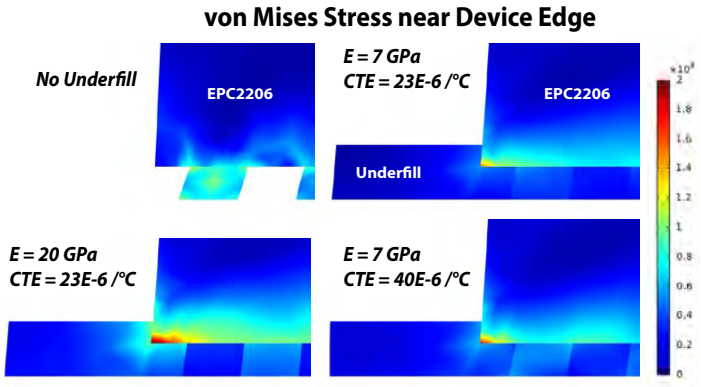


Figure 4-43 Von Mises (peak shear stress) in the edge-most solder bar under a temperature cycle change of ΔT = +100°C. Four different underfill conditions are simulated, with changing Youngs modulus (E) of the under- fill and different CTE as well. Note that deformation has been exaggerated by the same scale in each picture.

The effect of underfill on TC reliability was studied using EPC2218A [59] under the TC1 conditions of –40°C to 125°C, where two group of parts were compared: one with and one without underfill material. The underfill material selected was from Henkels Loctite (part number: Eccobond-UF 1173) which showed good performance in previous studies [60]. All parts were mounted on PCBs with two-copper-layer PCB using standard FR4 prepreg material. All underfilled devices were subjected to a plasma clean process prior to the underfill application. After every TC interval, electrical screening was performed. Exceeding datasheet limits was used as the criterion for failure. Physical cross-sectioning and SEM inspection were followed to further examine the electrical test failures. Solder joint cracking was found to be the primary failure mode throughout all failures analyzed. The experimental results from the test-to-fail approach are summarized in Weibull plots in Figure 4-44.

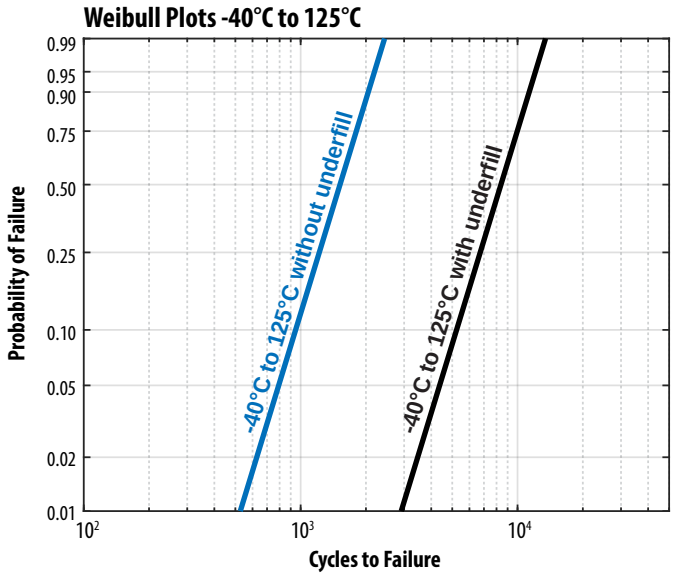


Figure 4-44. Weibull plots of temperature cycling (TC) results for EPC2218A, comparing with and without underfill material on a 2 Cu-layer PCB with a total PCB thickness of 1.6 mm.

The group without underfill reached more than 50% cumulative failures at 1600 cycles. The group with underfill showed no outlier devices were found in the measured $R_{DS(on)}$, nor in $R_{DS(on)}$ shift after 3000 cycles of TC1 stress. All parameters examined showed very tight distributions throughout all TC intervals. Physical cross-sectioning was conducted randomly on the 3000-cycle passing devices, where no solder joint cracking was observed. This shows that applying proper underfill material can significantly improve the TC capability. Therefore, the Weibull fit line in Figure 4-44 with the underfill leg is merely the lower bound confidence level based on the current test results.

This underfill TC study also forms the basis for the temperature-cycling reliability analysis of solar and DC-DC converters presented in Sections 5.1.6 and 5.2.5, respectively.

4.4.4. Development of a Lifetime Model for Power Cycling (PC) Stress

Power cycling (PC) test simulates the thermomechanical stress caused by the cyclic non-uniform thermal gradient between the packaged components and the PCB, due to the repetitive on and off operations of power devices. The PC test is complementary to the traditional temperature cycling (TC) test [JESD22-A104], in which the devices under test (DUTs) are usually non-operational and an insignificant temperature gradient exists between the PCB to the DUT. Figure 4-45 illustrates the different device junction temperature profile, comparing PC with TC.

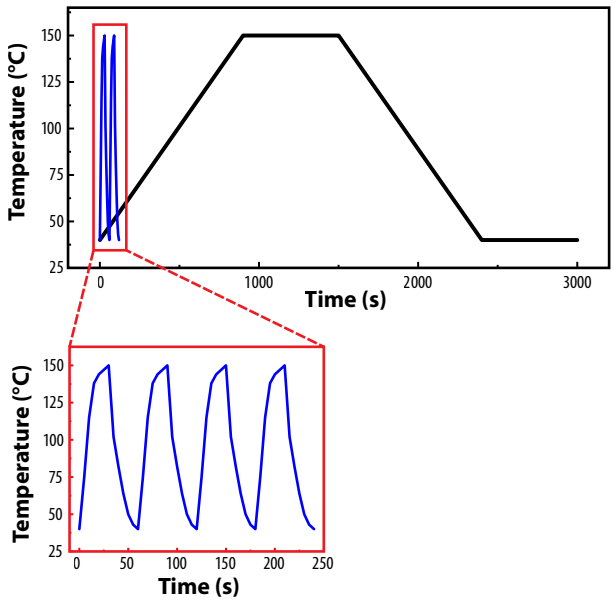


Figure 4-45. A comparison of device junction temperature between TC (50 minutes per cycle) and power cycle (60s per cycle).

A constant power approach [61] is used to heat up the DUTs, during which the junction temperature of the DUTs can reach the predefined maximum temperature (T_{Max}). After the heating cycle, the constant power is removed to initiate the cooling cycle, where external cooling fans are deployed to force airflow through the backside of

the DUTs. This enhanced cooling process enables the DUTs to reach the target minimum temperature (T_{Min}) within a specific time interval. The junction temperature of DUTs during PC is monitored by a FLIR® (Model T300) thermal camera and thermocouples that are mounted on the backside of the DUTs.

4.4.4.1. Applying the Thermomechanical Lifetime Model for Power Cycling Stress

In this section, the previously developed thermomechanical lifetime is found to be applicable to power cycling (PC), with some power exponents needing adjustment to be specific to PC stress. Therefore, a PC lifetime model was developed for EPC's PQFN devices under various PC stress conditions.

4.4.4.2. Test-to-fail: Power Cycling

The test-to-fail approach was adopted for PC testing of various EPC's PQFN devices with a fixed junction temperature variation from 40°C (T_{Min}) to 150°C (T_{Max}). The external bump layout of all PQFN DUTs investigated is identical, but the internal GaN device has two different dimensions, as shown in Figure 4-46. This work found that different internal die geometries significantly impacted the PC lifetime. Additionally, varying PC cycle time (t_{Cycle}) was found to cause significant differences in PC lifetime. Both effects were investigated systematically and later incorporated in the PC lifetime model presented in this report.

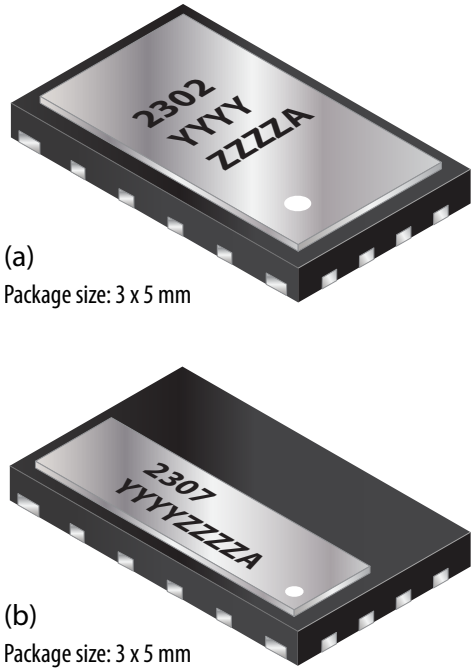


Figure 4-46: EPC's PQFN devices under test; (a) EPC2302, a Full-Die PQFN, (b) EPC2307, a Half-Die PQFN

After a certain number of power cycles, electrical parameters testing was performed, with failures determined by exceeding datasheet limits. The primary electrical failure characteristic is an increase in $R_{DS(on)}$. The Weibull plots from experimental results are summarized in Fig 4-47. The calculated Weibull MTTF along with the test conditions are tabulated in Table 4-8.

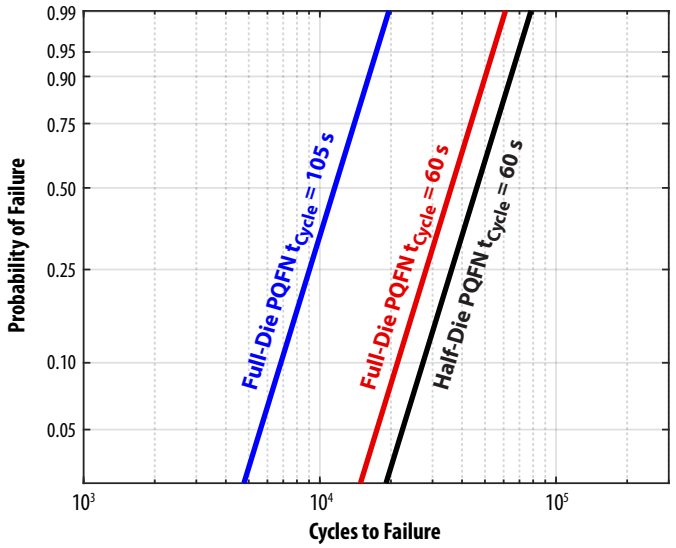


Figure 4-47. Weibull analysis of Full-Die and Half-Die PQFN packages tested for PC.

Product	ΔT (°C)	T_{Max} (°C)	T_{Min} (°C)	Cycle Time (s)	Cycle per day (*f* in Eq. 4-25)	MTTF (cycles)
Full-Die QFN (EPC2302)	110	150	40	60	1440	35,870
Half-Die QFN (EPC2307)	110	150	40	60	*1440*	46,020
Full-Die QFN (EPC2302)	110	150	40	105	*823*	11,519

Table 4-8: A summary of three PC Testing conditions and the respective MTTF determined by Weibull plots.

Physical cross-sectioning and SEM inspections were conducted to further examine the electrical screening failures. The corner gate bump solder joint cracking between the PQFN package and the PCB was identified as the primary cause of the PC failures, as shown in Fig 4-48. The failure mode observed in PC is similar to the TC failure modes, as discussed earlier.

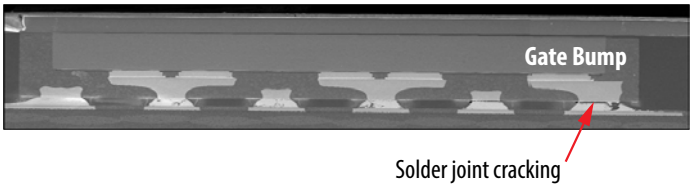


Figure 4-48. A representative SEM cross-section of a power cycle failure, highlighting the corner gate solder joint cracking is also responsible for the PC failures similar to TC.

The FA image of Figure 4-48 shows that the shear strain-type mechanical stress, caused by the CTE mismatch between the PCB and the QFN package, is primarily responsible for the PC failures. Therefore, the thermomechanical lifetime model in Eq. 4-25, developed from the Norris-Landzberg model, can be applied to PC stress. However, some fitting parameters, such as the power exponents, need to be adjusted to be specific to PC stress.

The Norris-Landzberg model, as implemented in Eq. 4-25 for TC, can be expressed in terms of shear strain observed by the solder joint to model the PC lifetime, as shown in Eq. 4-34 [44, 45].

$$N_{PC} \propto f^{\alpha} \cdot (\gamma_{ave}^{max})^{-\beta} \cdot exp\left(\frac{E_a}{kT_{Max}}\right) \quad \text{Eq. 4-34}$$

where f is the frequency term, representing number of the cycles per day, α is the frequency exponent, E_a is the activation energy that is most relevant at T_{Max}, the maximum temperature in the cycle, k is the Boltzmann constant; γ_{ave}^{max} is the average shear strain at corner gate bump between T_{Max} and T_{Min} and β is the power exponent. The shear strain γ_{ave}^{max} can be further defined by Eq. 4-35 [45].

$$\gamma_{ave}^{max} = DNP^{max} \cdot \Delta CTE \cdot \Delta T \cdot \frac{1}{h} \quad \text{Eq. 4-35}$$

where DNP^{max} is the distance from the maximum strain site, typically located at the extremities of the components, to the stress-neutral site, usually the device center. Since the half-die PQFN and full-die PQFN have identical dimensions, bump layout, and footprint, the DNP^{max} remains the same for all devices investigated in this study. ΔT is the difference between the high and low temperature extremes, and h is the standoff height (solder height) connecting the component and PCB. Similarly, the assembly and test conditions are consistent throughout the work, suggesting that the h and ΔT are also the same.

In the following section, three critical aspects involving PC stress and PQFN package implementation are investigated, as follows:

1. Effect of Cycle Time (t_{Cycle})
2. Effect of Internal Device Dimensions
3. Effect of ΔT (T_{Max} - T_{Min})

4.4.4.2.1. Effect of Cycle Time (t_{Cycle})

Two different cycle times (t_{Cycle}) were investigated for the full-die PQFN package with a fixed ΔT of 110°C. The device junction temperature profiles were measured using thermocouples directly mounted on the backside of the devices, with the measurements illustrated in Figure 4-49. Since the constant power testing method was used for the PC stress, little dwell time is observed at the T_{Max} or T_{Min} in both profiles, significantly different from the TC test.

Table 4-8 shows that the MTTF for the shorter cycle time (t_{Cycle} = 60°C) is approximately 3 times that of the longer cycle time (t_{Cycle} = 105°C). Based on Eq. 4-34, since the T_{Max} and T_{Min} remain the same for these two PC test legs, the PC lifetime (N_{PC}) can now be simplified to Eq. 4-36.

$$N_{PC} \propto f^{\alpha} \quad \text{Eq. 4-36}$$

where f represents the number of cycles per day, as listed in Table 4-8, and α is the power exponent. Therefore, α is estimated to be approximately 2. A positive frequency term exponent suggests that a shorter t_{Cycle} leads to a longer PC lifetime, and vice versa.

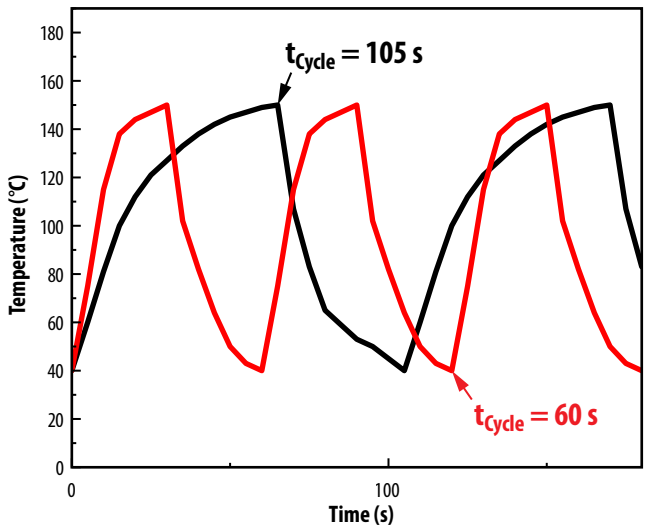


Fig 4-49: Comparison of a PC cycle temperature profile of 105 s per cycle and 60 s per cycle, where the temperature is measured by mounting a thermocouple directly on the backside of the PQFN packaged devices.

Within the heating period of a PC cycle, the device junction temperature quickly rises above the ambient temperature due to self-heating, while the temperatures of the PCB and the solder interconnects may lag behind. This contrasts with the traditional TC test, where the entire DUT assembly is placed in a TC chamber with a uniform temperature distribution. Therefore, a temperature gradient is expected between the device junction temperature and the PCB. The magnitude of the gradient is dependent on how quickly the device heats up to the target junction temperature (T_{Max}), which can be quantified by t_{Cycle}.

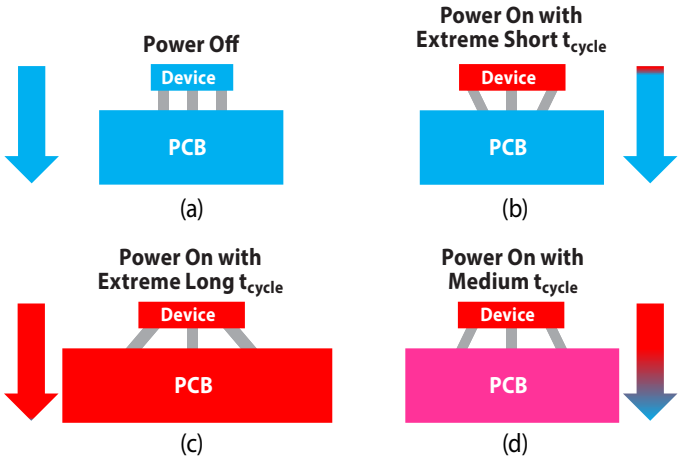


Figure 4-50: Comparison of thermomechanical strain observed by solder with condition (a): no power is applied, so both the device and PCB are blue; condition (b): the device is powered on with an extremely short t_{Cycle}, where the device is hot (red) but the PCB remains at ambient temperature (blue); condition (c) the device is powered on with an extremely long t_{Cycle}, where the device and PCB both reach T_{Max} (both are red); condition (d) the device is powered on with a medium t_{Cycle}, where the device is hot (red) but the PCB only reaches an intermediate temperature (pink).

The CTE of PCB is ~11.3 ppm/°C, significantly higher than the effective CTE of a full-die PQFN package device, which is estimated to be ~6.4 ppm/°C based on Table 4-9. Figure 4-50 (a) illustrates the neutral stress status when the power is off, where both the device and the PCB are at ambient temperature. Figure 4-50 (b) and (c) show the two corner scenarios under a PC cycle. In the case of Figure 4-50 (b), the heating interval is extremely fast, causing the PCB to lag completely and remain at the ambient temperature. In this case, the solder interconnect is expected to experience a small compressive shear strain at T_{Max} as the thermal expansion of the PCB is nearly zero, shown in Figure 4-50 (b). Figure 4-50 (c) illustrates the other extreme scenario with an extremely long t_{Cycle}, where the PCB temperature matches the device junction temperature, essentially making it a TC test without the dwell period. Figure 4-50 (d) demonstrates a medium t_{Cycle} case, where the PCB also heats up, but the temperature of the PCB remains significantly lower than the device junction temperature. In the cases of (c) and (d), the solder joints are expected to experience tensile-like shear strain at T_{Max} (junction temperature), with the magnitude of the shear strain depending on the t_{Cycle}.

Under normal PC testing, represented by the medium case in Figure 4-50 (d), a longer t_{Cycle} is projected to induce higher shear strain at T_{Max} in the corner gate solder joint, leading to a shorter PC lifetime (N_{PC}), and vice versa. This analysis validated the positive frequency power exponent of 2, which was calculated based on the test-to-fail data shown in Figure 4-47 and Table 4-8. Finite element analysis (FEA) is currently underway to further understand the underlying mechanism. Figure 4-51 shows the normalized PC lifetime as a function of varying t_{Cycle}, with all PC lifetimes normalized to the MTTF at a t_{Cycle} of 105 seconds. Users can refer to this plot to extrapolate the PC lifetime for different t_{Cycle}.

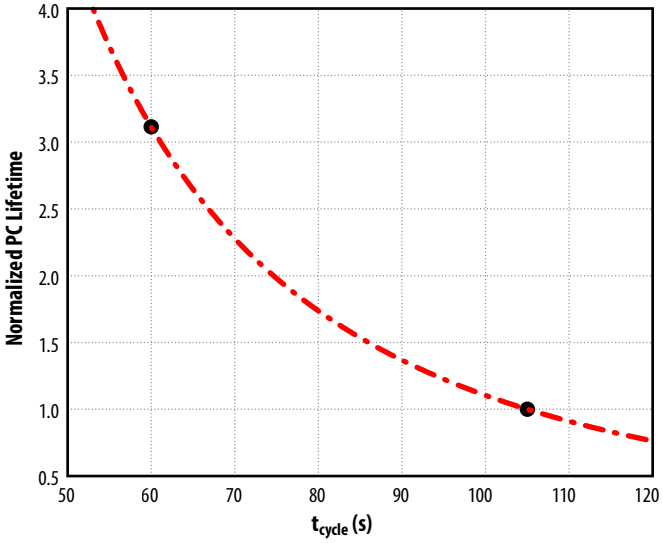


Figure 4-51: Normalized PC lifetime vs. cycle time (t_{Cycle}) with all PC lifetime normalized to the MTTF at a t_{Cycle} of 105 seconds.

4.4.4.2.2. Effect of Internal Device Dimensions

As illustrated in Figure 4-46, two PQFN-packaged devices, characterized as full-die and half-die QFN, have two different internal die dimensions and geometries. The Weibull distribution plot in Figure 4-47 and Table 4-8 shows that the half-die PQFN device (EPC2307) had a 30% longer lifetime than the full-die PQFN device (EPC2302), both using a t_{Cycle} of 60 seconds. In this section, the discussions will focus on explaining the lifetime difference between two PQFN package implementations.

Initially, this was a surprising result, as both PQFN-packaged devices had identical package dimensions, external pad layout and footprint. Additionally, the PCB assembly and PC test conditions were consistent throughout, suggesting h, f/t_{Cycle}, and ΔT of the device junction temperature are all the same. Therefore, the only difference between the two test legs that explains the variations in PC lifetime is the internal die size. Table 4-9 summarized the volumetric ratio of the lead frame (Cu), mold compound and the GaN-on-Si internal die of the full-die and half-die PQFN packages, from which the effective CTE of the two PQFN packaged devices were estimated. The calculation shows that the difference in the die/mold compound volumetric ratio within the PQFN packages can lead to significant variations in the effective CTE of the packaged devices.

	Lead Frame (CTE=17 ppm/°C)	Mold Compound (CTE=7 ppm/°C)	GaN-on-Si Die (CTE=2.6 ppm/°C)	$CTE_{Effective}$ (ppm/°C)
Full-Die QFN	14%	39%	47%	6.4
Half-Die QFN	15%	60%	25%	7.4

Table 4-9: Effective CTE estimates of the full-die and half-die PQFN packages based on the volumetric ratio of the lead-frame, mold compound, and the internal GaN-on-Si die.

In Eq. 4-35, the ΔCTE is the CTE difference between the PCB and the device. Thus, after combining Eq. 4-34 and 4-35, the PC lifetime (N_{PC}) can be modeled by Eq. 4-37, showing that the PC lifetime is inversely proportional to the CTE mismatch between the PCB and PQFN package.

$$N_{PC} \propto (\Delta CTE)^{-\beta} \qquad \text{Eq. 4-37}$$

where the power exponent β is estimated to be approximately 2 for SAC305 solder interconnects [50, 51, 52, 55, 56].

Table 4-8 shows that the MTTF of the full-die PQFN device is ~22% lower than that of the half-die PQFN device. After calculating the ΔCTE between the effective CTE of the PQFN package and the PCB (11.3 ppm/°C), the CTE mismatch of the full-die PQFN (ΔCTE = 4.9 ppm/°C) is 1.26 times higher than that of the half-die device (ΔCTE = 3.9 ppm/°C). Using Eq. 4-29 with a power exponent (β) of 2 [1,41], the calculated PC lifetime of the full-die PQFN device is projected to be ~37% lower than that of half-die PQFN device.

Although the estimated PC lifetime difference follows a consistent trend with the measurements, a small discrepancy was still observed. The first possible cause is that the power exponent of 2 was based on literature that studied the TC stress in SAC305 solder. Although Eq. 4-37 is still expected to hold, the actual power exponent for PC stress could be smaller than 2. In fact, using a β of 1.1 provides a good match between the projection and the experimental results, as β is usually reported in the range of 1 to 2 in literatures for TC stress [41,47]. Secondly, the discrepancy between the modeling and the experimental results is also likely due to the fundamental difference in heating methods between TC and PC stresses. Eq. 4-37 was derived from the original Norris-Landzberg TC model, which assumes that the temperature variations across the entire DUT/PCB assembly stack are uniform throughout the stress. Therefore, the ΔT between two different PQFN devices is identical, effectively canceling each other out. However, under normal PC stress, a thermal gradient between the device and the PCB is inevitable, as illustrated in Figure 4-50 (d). A slightly different thermal gradient profile from the device to the PCB can be expected when the heating source is a full-die compared to a half-die. FEA simulation work is underway to investigate this effect and further refine the PC lifetime model.

4.4.4.2.3. Effect of ΔT (ΔT = T_{Max} – T_{Min})

The effect of the device junction temperature difference (ΔT= T_{Max} –T_{Min}) within a PC cycle can also be modeled by Eq. 4-34, where the underlying wearout mechanism is identical to that of the TC test. Therefore, the PC lifetime model can be simplified to Eq. 4-38 by assuming the t_{Cycle} and T_{Max} remain the same.

$$N_{PC} \propto (\Delta T)^{-\beta} \cdot exp\left(\frac{E_a}{kT_{Max}}\right) \qquad \text{Eq. 4-38}$$

where the power exponent (β) of 2 is used for SAC305 solder [56] and the activation energy (Ea) is ~0.2 eV based on previous TC testing results for SAC305 [1]. Additional testing is ongoing to further refine them, making them more applicable to PC stress.

Figure 4-52 shows the PC lifetime of the full-die and half-die PQFN packages at 50% failure rate and 1% failure rate with t_{Cycle} = 60 seconds and T_{Max} = 150°C. The PC lifetime increases significantly as the ΔT decreases, due to the reduced thermomechanical stress and strain in the solder joints, consistent with the observations from TC stress.

Figure 4-53 shows the projected PC lifetime with a 1% failure rate for full-die PQFN and half-die PQFN devices with four different T_{Max} values: 50°C, 75°C, 100°C, and 125°C. Figure 4-53 highlights the significance of the T_{Max} term, which suggests that higher T_{Max} is expected to induce more solder creep failure, leading to shorted a PC lifetime.

4.4.5. Conclusion

A comprehensive lifetime model, based on the experimental power cycling (PC) data for EPC's PQFN packaged devices, was developed with the understanding that corner gate solder joint cracking is the primary cause of PC failures. The PC-specific lifetime model considers the effects of cycle time (t_{Cycle}), ΔT (T_{Max} – T_{Min}), and the internal die dimensions and geometries.

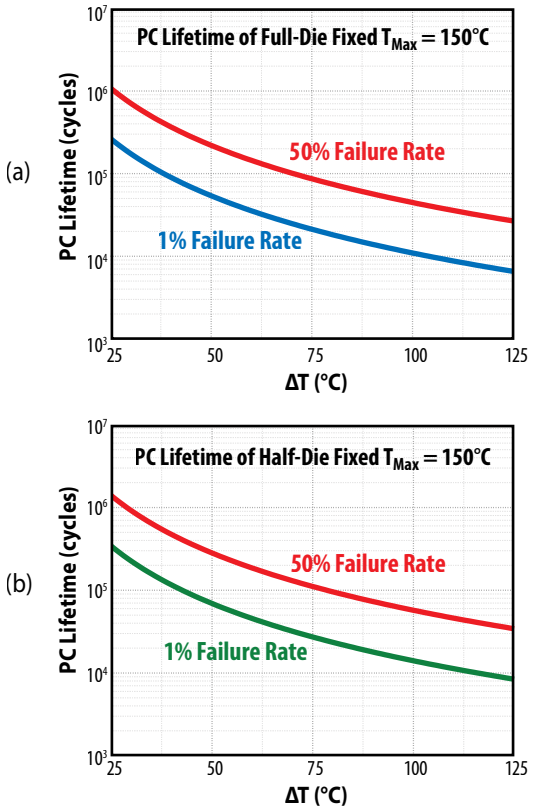


Figure 4-52: PC Lifetime prediction curve with respect to ΔT for a varying T_{Min} and a fixed T_{Max} of 150°C with a) Full-Die PQFN; b) Half-Die PQFN.

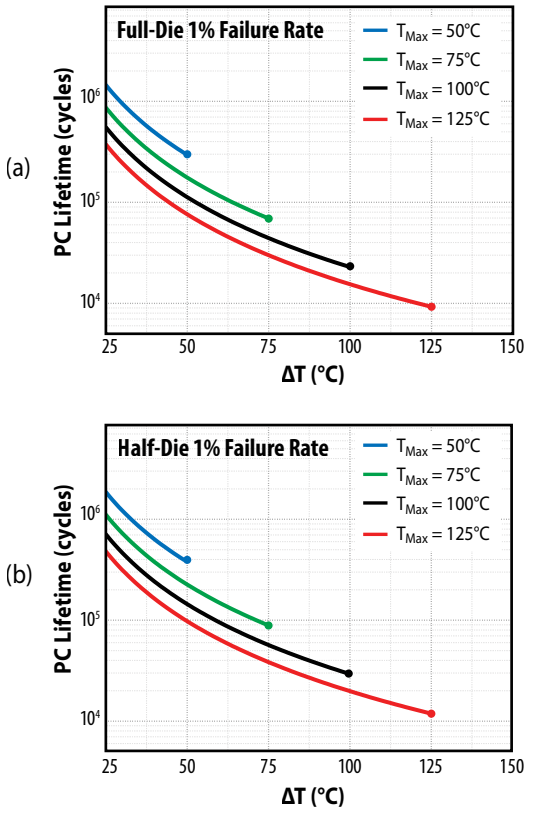


Figure 4-53: Lifetime curves for 1% failure rate vs. ΔT at T_{Max} of 50°C (blue), 75°C (green), 100°C (black), and 125°C (red) of a) Full-Die PQFN; b) Half-Die PQFN.

4.5. Mechanical Stress Wear-Out

4.5.1. Introduction to Mechanical Stress Wear-out Mechanisms

The lifetime of a product, or its suitability in applications, may be limited by the mechanical stresses encountered. In this section, some of the most common mechanical stressors, die shear, backside pressure, and bending force are characterized. The CSP and QFN package are demonstrated to be robust under normal assembly or mounting conditions.

4.5.2. Die Shear Test of Chip-Scale Parts

The purpose of die shear test is to evaluate the integrity of the solder joints used to attach eGaN devices to PCBs. This determination is based on the in-plane force at which, when applied to a mounted device, the die shears from the PCB. All testing followed the military test standard, MIL-STD-883E, Method 2019 [62].

Figure 4-54 shows the test results of four selected GaN transistors. Ten parts were tested for each product. The smallest die tested is EPC2036/EPC2203, which only has four solder balls with a diameter of 200 µm and a die area of 0.81 mm^2. As expected, this product turned out to have the least shear strength, however, it exceeds the minimum force requirement specified by the MIL standard, as shown in Figure 4-54. The largest die tested was EPC2206, a land grid array (LGA) product with die area of 13.94 mm^2. EPC2206 exceeds the minimum force requirement by more than a factor of ten. Within the size spectrum, two additional products were tested: EPC2212 (100 V LGA) and EPC2034C (200 V BGA). Both products surpassed the minimum force significantly. Figure 4-54 shows that all WLCSP GaN products are mechanically robust against environmental shear stress under the most stringent conditions.

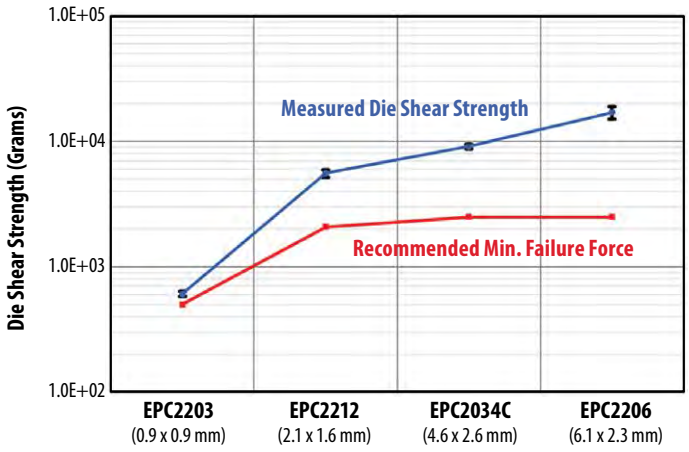


Figure 4-54: Various die sizes and solder configurations of GaN transistors were tested to failure while measuring the shear strength. The results are shown with black dots. The red dots show the minimum recommended die shear strength under MIL-STD-883E, Method 2019.

4.5.3. Backside Pressure Test of Chip-Scale Parts

Another critical aspect of the mechanical robustness of GaN devices is how well they handle backside pressure. This is an important consideration for applications that require backside heatsinking to die. It is also important to determine the safe "pick-and-place" place force during assembly.

Backside pressure tests up to 400 psi were performed, where the pressure is calculated by the force applied divided by the die area. The pressure was applied directly to the backside of the die using a loading speed of 0.6 mm/min. Before and after the pressure test, parametric testing was performed to determine pass or fail. Subsequently, the parts were exposed to humidity bias testing (H3TRB) at 60 V_{DS}, 85°C, and 85% relative humidity for 300 hours. H3TRB is effective to determine if there were any latent failures caused by mechanical damage (internal cracking) from the pressure test.

EPC2212 (100 V, LGA) and EPC2034C (200 V, BGA) were tested, and both passed 400 psi. The 400 psi is calculated by normalizing the force applied on the backside of the device (Si substrate) to the die area. Results show that GaN transistors have enough margin to handle backside pressure that is normally used at a PCB assembly house. Though these parts survived 400 psi, backside pressure should be limited to 50 psi or less.

4.5.4. Bending Force Test of Chip-Scale Parts

The purpose of the bending force test is to determine the ability of a GaN transistor to withstand flexure of the PCB, which might occur during handling, assembly, or operation. Though this test standard was developed for passive surface mount components (AEC-Q200) [63], many customers have concerns about bending forces on GaN transistors for two main reasons:

1. Robustness of the WLCSP solder joints
2. Piezoelectric effects within the transistor that may alter device parametric values and disrupt circuit operation

To address these concerns, bending force testing on four EPC2206 devices following the AEC-Q200-005A test standard [64] were conducted. Devices are assembled near the center of an FR4 PCB (100 mm long x 40 mm wide x 1.6 mm thick). With ends rigidly clamped, a force is applied on the opposite side from the device, leading to an upward deflection of the PCB. After a 60 second dwell in this flexed state, all device electrical parameters are measured.

Table 4-10 shows normalized $R_{DS(on)}$ versus board deflection for all four devices under test. All devices passed the 2-mm test requirement. Two devices failed at 6-mm deflection, while the remaining two survived all the way to 8 mm. Postmortem analysis revealed that the failure mode was solder joint cracking, leading to an open gate connection. Up until failure, $R_{DS(on)}$ did not show any appreciable response to board flexure. The same result was observed in other electrical characteristics like V_{TH} and I_{DSS}.

	0 mm	2 mm	4 mm	6 mm	8 mm
DUT1	1.00	1.01	1.00	0.98	0.98
DUT2	1.00	1.02	1.01	*Failed*	-
DUT3	1.00	1.01	1.03	*Failed*	-
DUT4	1.00	0.99	0.99	1.03	1.04

Table 4-10: Normalized $R_{DS(on)}$ versus board deflection for four devices during bending force test.

Note: Values are normalized to $R_{DS(on)}$ in the unflexed case. Two of four devices failed at 6-mm deflection, while the remaining two devices survived 8 mm. No significant stress response was seen in any device parameter.

4.5.5. Bending test on PQFN devices

PCB bending test was conducted to evaluate the solder joint robustness between the power quad-flat no-leads (PQFN) package devices and PCB under PCB bending and warpage stress conditions. These tests will address the customers' concerns in the module assembly, handling, and operations when potential mechanical impacts are present, such as PCB deformation in the motor drive applications and mechanical shock and PCB bending in automotive-related applications. The bending test uses a 3-point bending setup, following the Substrate Bending Test as described in IEC-60068-2-21. The devices are assembled at the center of an 8-layer PCB with the size of 180 mm long, 90 mm wide, and 1.6 mm thick. The PCB is placed on two supporting fixtures with a 90 mm gap. The device under test is placed facing down. The bending tool applies the force downwards at the back of the PCB to force the bending deflection. The test setup is shown in Figure 4-55.

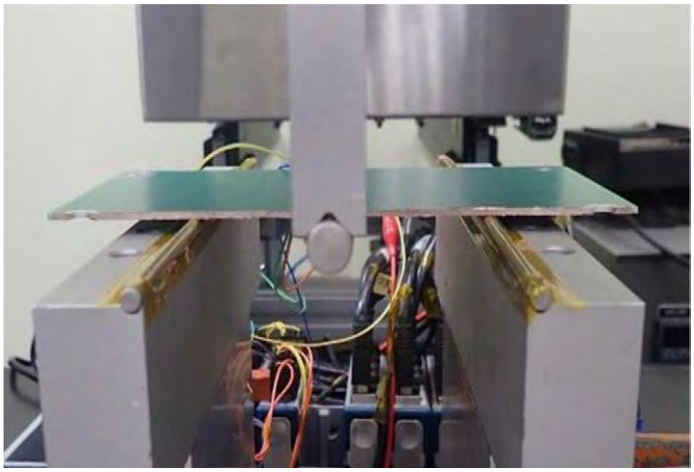


Figure 4-55 Setup of the bending test used for evaluating PQFN devices.

Daisy-chain PQFN devices are used to enable reliable in-situ monitoring of the solder joint resistance during the test. The daisy-chain PQFN devices are developed and manufactured using the same PQFN component layout, constructions, and materials as the EPC2302.

The first test condition evaluates the solder joint robustness under constant load. 10 devices were stressed up to 2mm bending deflection over 20s duration. The resistance of the daisy-chain devices was in-situ monitored during the bending test. Table 4-11 shows the resistance of each device before and after the test. For all the 10 devices, the resistance change is minimal, suggesting that no degradation in the solder joint is generated from this test. To further verify the solder joint quality, three devices were randomly picked for solder joint cross-section inspection, which showed no observable solder joint cracks in the cross-sections, agreeing with the resistance records. Thus, these results show that the PQFN solder joints can handle constant load from PCB bending with a high level of reliability.

A second test condition evaluated robustness versus PCB bending in a test-to-fail manner. The purpose of test-to-fail is to understand the ultimate failure mode under extreme stress conditions that are well beyond the normal operating conditions. 10 devices were tested. The bending deflection gradually increased from zero up to max 15 mm, or when an abrupt resistance change happens. The bending deflection step is 1 mm and the test duration at each step is 20 s.

Table 4-11 shows the resistance record before and after the bending test, and the max bending deflection. All the devices passed the bending deflection up to 10 mm, where insignificant amount of resistance change was observed. Two devices failed at approximately 11 mm of bending deflection. Failure analysis was conducted on the two failure devices and revealed that the failure mode is cracks in the GaN-on-Si die. Solder joint cross-sections were conducted on the failure die, which did not show observable solder joint cracks. Thus, the PQFN devices can survive PCB bending up to 10 mm, without observable degradations in the solder joints.

Item	Sample No.	Pre-Test Resistance (Ω)	Post-Test Resistance (Ω)	Max Deflection (mm)
Condition 2	1	0.27	0.26	15.00
	2	0.26	1.78	11.36
	3	0.24	0.24	15.00
	4	0.23	0.23	15.00
	5	0.26	0.26	15.00
	6	0.23	0.22	15.00
	7	0.22	0.23	15.00
	8	0.21	0.22	15.00
	9	0.23	0.23	15.00
	10	0.23	0.86	10.82

Table 4-11: Resistance before and after the second test, and the max bending deflection

SECTION 5.0 MISSION-SPECIFIC RELIABILITY PREDICTIONS

Section 5 introduces a framework for analyzing device lifetimes in mission specific applications with complex stress conditions and varying durations. The analysis is primarily based on the mathematical development presented in Section 3 of the reliability report.

To perform a mission-specific reliability prediction, the process typically consists of two steps. The first step involves identifying the main stressors most relevant to the mission application. Once this step is established and the individual lifetimes are estimated, Eq. 5-1 can be used to project the overall system-level reliability.

$$\frac{1}{MTTF_{Total}} = \frac{1}{MTTF_1} + \frac{1}{MTTF_2} + \cdots + \frac{1}{MTTF_i} \quad \text{Eq. 5-1}$$

where 1, 2, …, and i correspond to the individual stressors relevant to the specific mission and application, with each stressor present throughout the mission.

However, the devices typically do not operate continuously under a single condition, but rather under quite complex conditions, with varying voltage, temperature, or frequency during different periods throughout the mission lifetime. Therefore, the second step is to further examine the operating conditions and stress profiles within each stressor's mission lifetime. One good example is solar panels placed outdoors, which experience varying temperature profiles throughout the years. Different temperature profiles can lead to significantly different temperature cycling (TC) lifetime, suggesting that using a single or averaged temperature varying profile is very unlikely to accurately predict the thermomechanical reliability of the solar panel. Therefore, Eq. 5-2 was developed to address these concerns.

$$\frac{1}{LT_{Total}} = \frac{a}{LT_a} + \frac{b}{LT_b} + \cdots + \frac{n}{LT_n} \quad \text{Eq. 5-2}$$

where a, b, …, n in the numerators represent the fractional operation time of each individual stress condition and LT_n is the respective lifetime under each stress condition.

In this section, this framework is applied to three example applications: solar, DC-DC, and lidar.

5.1. Solar Application Specific Reliability

5.1.1. Introduction

Microinverters and power optimizers are widely utilized in modern solar panels to maximize energy efficiency and conversion. Such topologies and implementations usually require a minimum of 25 years of lifetime, which is becoming a critical challenge for market adoption. Low-voltage gallium nitride (GaN) power devices (V_{DS} rating < 200 V) are a promising solution and are being used extensively by an increasing number of solar manufacturers.

In this section, a test-to-fail approach is adopted and applied to investigate the intrinsic underlying wear-out mechanisms of GaN transistors. The study enables the development of physics-based lifetime models that can accurately project the lifetimes under the unique demands of various mission profiles in solar applications.

5.1.2. Trends In Photovoltaic Power Conversion

The ever-increasing demand for renewable energy sources has led to a rapid growth in rooftop solar installations across residential and commercial sectors. Traditionally, string inverters have been widely employed in solar installations, where multiple solar panels are connected in series. The inverter is responsible for converting direct current (DC) output from solar panels to alternating current (AC) electricity that can be used to power homes.

String inverters have served as a reliable choice for years. However, they also face many challenges, including reduced performance due to shading, panel mismatch issues, and a lack of module-level monitoring. Most importantly, due to the series configuration of the string inverters, the lowest performing panel dominates the energy conversion rate of the entire system, which could significantly lower the system efficiency.

The Department of Energy released the $1/watt photovoltaic (PV) system initiative in 2010, where developing higher efficiency and more reliable module-level integrated inverters was highlighted as the key area of improvement to meet the target [65]. The SunShot 2030 PV program envisions a similar cost target by 2030 [66]. To meet the goals and maximize energy production, emerging technologies such as microinverters and power optimizers have gained significant attention.

Microinverters are small, individual inverters that are attached to each solar panel, allowing for DC to ac power conversion at the panel level. This enables each solar panel to function at its peak performance by using independent maximum power point tracking (MPPT). Even if a tree branch shades certain panels, all the neighboring panels can still convert at their full capacity. The drop in efficiency only affects the panels in the shade.

Independent tracking also allows solar users to monitor the health of each panel easily. If a panel requires repair, it won't bring down the whole system. In addition, microinverters make it easy to add panels to increase power output. Microinverters can be more expensive than string inverters but can pay off over time by getting more power from your system. Therefore, microinverters in the market need to match panel guarantees with 25-year warranties [67,68].

Power optimizers are DC-DC converters integrated into the solar panel wiring, enabling MPPT of each individual solar panel by continually regulating the DC characteristics to maximize energy output. A power optimizer is a good solution for situations where shading is an issue, or the panels must be placed on multiple roof surfaces with different orientations. Therefore, power optimizers generally are a more energy efficient solution than string inverters. The power optimizer also requires 25 years of warranty [69,70].

5.1.3. Applying Test-to-Fail for Solar

After reviewing the benefits that are driving the switch from string inverters to microinverters and power optimizers in photovoltaic systems, the test-to-fail methodology is introduced and the three device "stressors" most likely responsible for device failure are identified—gate bias, drain bias and temperature cycling. In the subsequent sections, the impact of each of these factors on device lifetime, expressed in terms of mean time to failure (MTTF) and other parameters, is assessed.

To address the reliability concerns surrounding the requirement for 25 years of reliable operation, a test-to-fail approach [4,27] is adopted and applied to GaN devices that are commonly used in

solar applications. The methodology involves stressing the devices under test (DUTs) to cause them to fail quickly under accelerated conditions while monitoring type and time of failure.

By analyzing the failures and understanding the underlying failure mechanisms, physics-based lifetime models can be developed to explain the unique characteristics of GaN. The developed models can be used to accurately project the lifetimes under all mission profiles that are unique to solar applications.

By examining the mission profiles for solar applications, three key reliability stressors are identified; gate bias, drain bias and temperature cycling (TC). The total MTTF can be described by Equation 5-3.

$$\frac{1}{MTTF_{Total}} = \frac{1}{MTTF_{Gate}} + \frac{1}{MTTF_{Drain}} + \frac{1}{MTTF_{TC}} \qquad \text{Eq. 5-3}$$

Therefore, it is critical to understand which stressor is the limiting factor in reliability. This stressor warrants more consideration during design and operation. Each stressor is studied independently by using this test-to-fail approach, where the individual intrinsic wear-out mechanism is successfully identified, and the corresponding lifetime is determined.

5.1.4. Gate Bias

GaN high electron mobility transistors (HEMTs) are used in DC-AC (microinverters) or DC-DC (power optimizers) topologies in their solar applications. The gate terminal must be biased periodically during switching. Hence, gate reliability over time is the first stressor to examine. As shown in Figure 4-2 (Section 4.1.2), GaN HEMTs have an approximately 1-ppm failure rate projected after 25 years of continuous dc bias at $V_{GS(max)}$ = 6 V.

5.1.5. Drain Bias

The low on-resistance ($R_{DS(on)}$) and small die size of GaN HEMTs significantly increase the power conversion efficiency and reduce the power losses in microinverter and DC-DC converter applications. However, one common concern for GaN is dynamic on-resistance.

The flyback is one of the more popular topologies for microinverters in solar applications. When selecting the appropriate GaN transistors for the primary side, three main contributing factors to the drain voltage are considered. These are (1) the bus voltage, (2) the flyback voltage, and (3) the voltage overshoot due to ringing caused by the parasitic inductance in the design. The typical bus voltage for a microinverter is 60 V in a solar application. The flyback voltage is determined by the product of the system's output voltage and the turns ratio of the transformer. By adding some margin for the voltage overshoot and derating, a 170-V maximum V_{DS} rating is frequently desired by the solar customers using such topology.

The EPC2059 [71] is a 170-V maximum V_{DS} rated product that meets the general requirements for microinverters in solar applications. Figure 5-1 shows the in-situ $R_{DS(on)}$ test results of a representative EPC2059 device that was operated under continuous hard switching at 136 V (80% of the max rated drain bias of 170 V) while the case temperature was modulated at 80°C. This temperature is used because it is considered the nominal operating temperature for solar panels. As shown in Figure 5-1, the lifetime model is plotted against the measured data. The model predicts the $R_{DS(on)}$ increase due to continuous hard switching in 25 years to be approximately 10%.

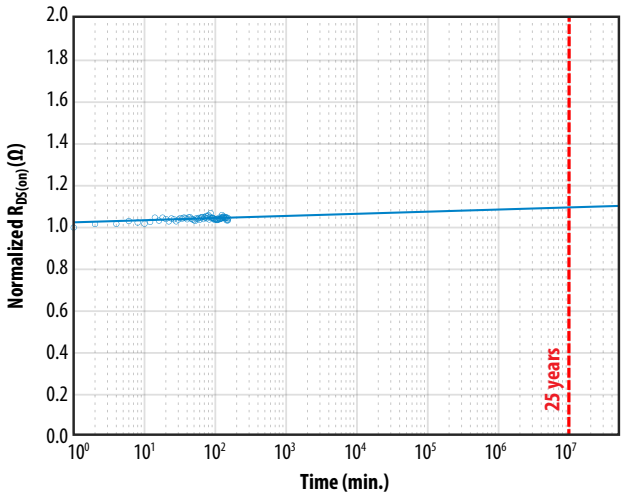


Figure 5-1. The projected $R_{DS(on)}$ shift of the EPC2059, a 170-V rated device, in 25 years of 100-kHz continuous hard-switching operation at 136 V is approximately 10%. The blue circles represent measured data.

Another popular option for solar systems is to use a DC-DC converter in a power optimizer. This has been adopted by many solar providers due to its superior efficiency. EPC's GaN devices such as the 100-V rated EPC2218 [72] and EPC2302 [73] among others, are good fits for this application.

Figure 5-2 plots the results obtained with the lifetime model alongside the in-situ measured data for two representative devices—the EPC2218 and EPC2302. A shift of less than 10% in 25 years of continuous hard switching at 80% of the max rated drain bias and 100 kHz is expected. This result suggests that dynamic $R_{DS(on)}$ failure is not the dominant factor determining the lifetime for EPC's GaN devices in solar applications.

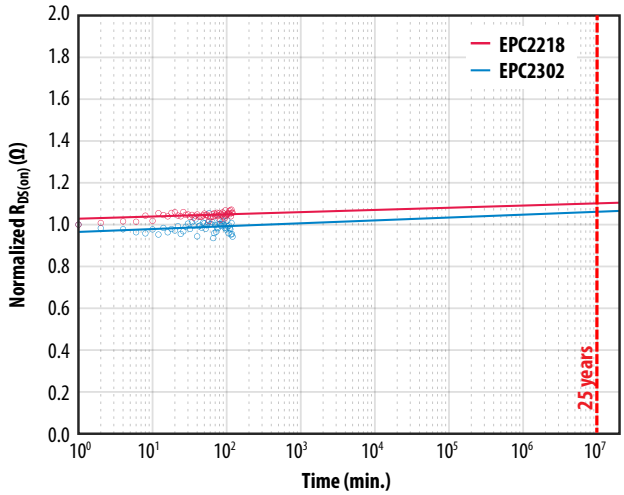


Figure 5-2. The projected $R_{DS(on)}$ shifts of the EPC2218 and the EPC2302, which both are 100-V rated devices, under continuous hard-switching operation at 80 V, 100 kHz are plotted here. The blue and red circles represent measured data.

5.1.6. Temperature Cycling

Temperature cycling is another critical area of particular interest for solar applications. Solar panels are placed outside, and experience significant ambient temperature changes each day. Therefore, devices mounted on the PCBs in the solar panels must be capable of surviving 25 years of continuous ambient temperature change.

In real world applications, solar panels experience varying ambient temperatures, and the amount of temperature change varies significantly depending on the season and location. As a result, a more-general lifetime model for thermo-mechanical stress is warranted to account for all mission profiles over the 25 years of lifetime. Another TC lifetime model is developed below to account for different ΔT at different seasons of the year, as shown in Equation 5-4.

$$\frac{1}{N_{Total}} = \frac{a}{N_{\Delta T_a}} + \frac{b}{N_{\Delta T_b}} + \cdots + \frac{i}{N_{\Delta T_i}} \qquad \text{Eq. 5-4}$$

where N_{Total} is the total calculated lifetime number of cycles, $N_{\Delta T_a}$ corresponds to cycles-to-failure for the condition of ΔT_a and a is the fraction of time the device was operational under the condition of ΔT_a, $N_{\Delta T_b}$ corresponds to cycles-to-failure for the condition of ΔT_b and b is the fraction of time the device was operational under ΔT_b, and $N_{\Delta T_i}$ corresponds to cycles-to-failure for the condition of $N_{\Delta T_i}$ and i is the fraction of time the device was operational under $N_{\Delta T_i}$.

There are three main factors that predominantly determine the lifetime of the solder joints:

1. The duration of each mission profile needs to be separated. This effect is accounted for by the fractional coefficient in the numerator of each term in equation (5-4), such as a, b, …, and i.
2. The temperature change (ΔT) in each mission profile. This term is addressed by the Norris-Landzberg model plotted in Figure 5-3. The solder joints experience the most stress during the period when the devices are subjected to the largest ΔT, which translates to the shortest cycles-to-failure. The overall lifetime of the device essentially will be dominated by the most stressful period. This effect is addressed by putting the cycles-to-failure terms ($N_{\Delta T}$) in the denominator and then summing them up collectively.
3. The hottest temperature extreme of each cycle, or the baseline temperature. For instance, the solder joints may experience different stress levels given an identical ΔT in the winter or in the summer.

Each of these factors is included in the analysis that follows, which is based on the board-level thermomechanical reliability study presented in Section 4.4.4, assuming a 0.1% failure rate for the EPC2218A with underfill.

The projected lifetime curves using the Norris-Landzberg model are plotted in Figure 5-3 assuming T_{Max} is 125°C, which is the worst-case scenario for the creep failure mechanism. The horizontal, black-dashed line at 9,125 cycles represents a duration of 25 years of continuous operation assuming one thermal cycle per day.

Figure 5-3 shows that after 25 years of continuous operation under a constant temperature swing of 72°C from hot to cold, or vice versa, only 0.1% of the EPC2218A devices with underfill material would fail the datasheet limit due to an increase in $R_{DS(on)}$ value. At a 1% failure rate, 99% of the devices should be capable of surviving 25 years of continuous operation when subjected to a constant ΔT of 95°C. Even without underfill material, 99% of the parts should survive a fixed ΔT of approximately 51°C over 25 years of lifetime.

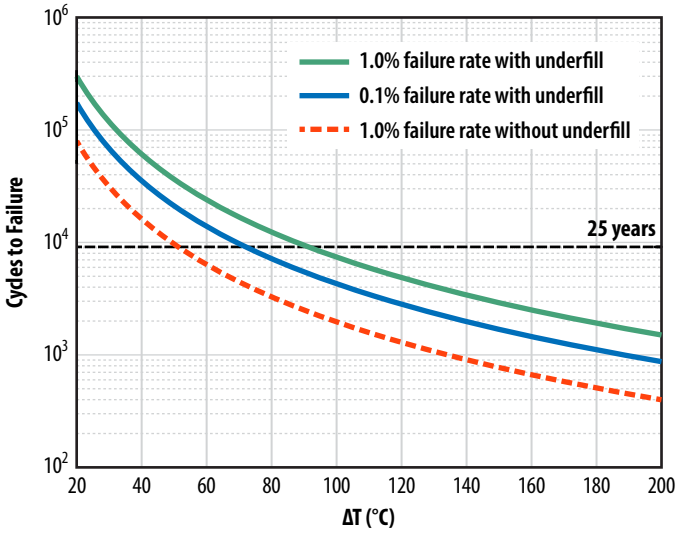


Figure 5-3. Lifetime prediction curves for EPC2218A with respect to ΔT using the Norris-Landzberg model.

Now let us examine a real-world example based on the lifetime model of Eq. 4-25. Assume that solar panel system is installed outdoors near solar panels in Phoenix, Arizona, U.S.A., where the climate is well-suited for solar, but also demands extreme temperature changes over time. Use the weather report history of Phoenix, Arizona as an example [74].

In addition, 30°C of device self-heating is added to the ambient temperature change for the total lifetime calculations. For the 0.01% failure rate, or 100 ppm, which means 100 devices failed in 1 million parts tested, the EPC2218A with underfill is projected to have 18,218 cycles to failure, equivalent to 49.9 years of lifetime operation considering one cycle per day for GaN devices in the example application.

If we extrapolate to a 0.001% failure rate, or 10 ppm, suggesting only 10 failures out of 1 million devices tested, now the total lifetime is calculated to be 10,971 cycles. This is equivalent to approximately 30 years of continuous operation with one cycle per day.

The results imply that temperature cycling is the most critical stressor that could be limiting the overall lifetime for GaN used in solar applications. However, by using proper underfill materials TC reliability can be significantly improved to exceed the required 25 years of continuous operation with a low failure rate under nominal solar mission profiles.

5.1.7. Conclusions

The test-to-fail results and physics-based lifetime projections show that neither gate bias nor drain bias are major reliability concerns for microinverters or power optimizers in solar applications. Using appropriate underfill materials can vastly reduce thermal cycling reliability risk, resulting in lifetimes exceeding 25 years.

5.2. DC-DC Application Specific Reliability

5.2.1. Introduction

DC-DC converters exist in virtually every application of modern power electronics. Due to small die size, low on-resistance, and low parasitic capacitance, GaN power devices offer superior conversion efficiency and record-setting power density. In this paper, test-to-fail methodology is adopted to investigate the intrinsic wear-out mechanisms such as would be experienced in common DC-DC converters. Devices are stressed under gate bias, drain bias, and temperature cycling individually. The lifetime of each stressor is therefore projected based on the physics-based model developed from test-to-fail and an understanding of the unique stress conditions in DC-DC converters.

GaN devices have demonstrated better switching performance and power density with figures of merit (FOM) 3 to 10 times superior to comparable silicon devices. This trend will only accelerate as GaN FETs continue to improve while Si MOSFET are already very close to their theoretical limits.

GaN devices have enabled easy to use topologies like the synchronous buck converter to reach new levels of efficiency and power densities. Taking advantage of reduced switching losses and no reverse recovery, designers can increase switching frequencies while also reducing power losses. This increase in switching frequency allows for smaller, more efficient inductors that in turn can increase efficiencies by further lowering resistive losses while reducing overall volume. The amount of capacitance can also be cost reduced and with better transient response. Overall, this leads to designs with higher power density, higher efficiency, and lower system cost, hence the broad adoption trends seen throughout various end markets.

GaN HEMTs are particularly valuable where power density is the goal. For example, designers have taken advantage of EPC wafer level chip scale packaging (WLCSP) to significantly increase the power density of intermediate bus converters (IBC) for server applications migrating to a 48V distribution rail. Many designers have chosen an LLC topology operated as DC transformer (DCX) with GaN in both primary and secondary sides. On the primary side the small size of GaN allows the devices to reduce conduction and gate drive losses in the same footprint as a power MOSFET, while the small COSS allows the LLC to operate with a higher power delivery cycle and better transformer utilization. On the secondary side GaN enables the lowest conduction losses in a given area while minimizing gate drive losses thanks to the very small QG. This combination of best-in-class power devices and advanced packaging technologies has allowed for record power densities [75].

5.2.2. Test-to-Fail Methodology

To address all the reliability concerns in common DC-DC converters, a test-to-fail methodology [1, 15, 4, 27,76] is adopted and applied to popular GaN devices. In DC-DC applications, three key stressors are identified; gate bias, drain bias and temperature cycling (TC). The total MTTF can be described by Equation 5-5,

$$\frac{1}{MTTF_{Total}} = \frac{1}{MTTF_{Gate}} + \frac{1}{MTTF_{Drain}} + \frac{1}{MTTF_{TC}} \qquad \text{Eq. 5-5}$$

5.2.3. Gate Bias

In DC-DC converters, the gate terminal of GaN HEMTs must be biased periodically during switching. GaN HEMTs have approximately 1 ppm failure rate projected after 25 years of continuous DC bias at $V_{GS(max)}$ = 6 V. This shows that gate bias stress is not the dominant stressor limiting the overall lifetime.

5.2.4. Drain Bias

A frequently discussed reliability concern for GaN under drain bias is dynamic on-resistance. This is a wear out mechanism where the $R_{DS(on)}$ of GaN HEMTs rises when the devices are subjected to high drain-source voltage (V_{DS}). One of the dominant mechanisms responsible for the increase in $R_{DS(on)}$ is hot electron induced trapping effects [1, 15, 4, 27, 76]. As the trapped charges accumulate, electrons from the 2DEG are depleted, leading to an increase in $R_{DS(on)}$. The detailed lifetime model derivation is discussed in Section 4.2.

The next sections address the following knowledge gaps:

1. How can a representative drain voltage waveform of a common DC-DC converter be correlated with various reliability testing topologies (stressors)?
2. What are the projected lifetimes of each individual reliability testing topology (stressor) based on the lifetime model developed from the electron trapping effect?
3. How does individual reliability lifetime prediction determine the overall lifetime of GaN devices?

First, a SPICE simulation was conducted for a buck converter using an EPC9078 demonstration board featuring 100 V EPC2045 GaN transistors [77]. To include the corner conditions for a real-world application, an intentionally poorly designed buck converter was simulated, where abnormally high parasitic inductances were added to emulate a worst-case scenario. Figure 5-4(a) shows the simulated turn-off voltage waveform, where the drain voltage immediately rings to a peak voltage of approximately 120 V and then the amplitude of ringing drops off quickly to stabilize at a bus voltage of 80 V. The simulated voltage waveform in Figure 5-4(a) can be deconvoluted by two separate voltage waveforms as shown in Figure 5-4(b) and (c). Figure 5-4(b) illustrates that the overvoltage ringing can be fitted with a set of half-sinusoidal voltage waveforms. After the ringing subdued and reaches the bus voltage,

the equilibrium part of the waveform can be modeled by a voltage waveform as shown in Figure 5-4(c). Waveforms in Figure 5-4(b) and (c) can be realized by two different reliability testing circuits, which will be discussed separately in the following discussions.

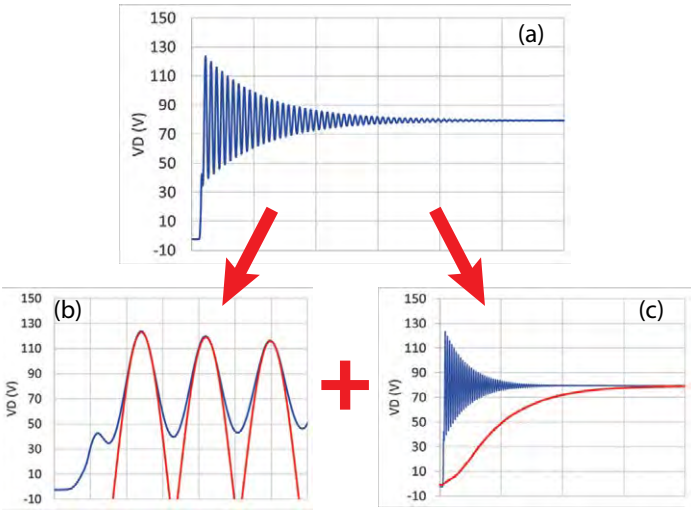


Figure 5-4 (a) A simulated turn-off drain voltage waveform based on a poorly designed buck converter, where a 120 V ringing and 80 V bus voltage are shown. (b) Ringing can be fitted with a set of half-sinusoidal waveforms. (c) The equilibrium portion of the waveform can be fitted by a different voltage waveform shown in red.

Transient overvoltage ringing is commonly observed in GaN HEMTs under high dV/dt switching conditions. Because GaN HEMTs lack avalanche mechanisms, the reliability impact under such transient overvoltage stress is becoming a critical challenge for the industry. To properly address this concern, an unclamped inductive switching (UIS) test circuit was developed as shown in Figure 5-5(a). Figure 5-5(b) shows a half-sinusoidal voltage waveform with a 120 V overvoltage spike that is generated by the UIS test system developed. This transient overvoltage testing was performed at 100 kHz repetitively with a 6% duty cycle during which the GaN HEMT is turned on and $R_{DS(on)}$ is monitored *in-situ*.

Figure 5-6 (a) shows in-situ measured $R_{DS(on)}$ of three representative EPC2218 devices [72] (100 V rated $V_{DS,Max}$) from three different manufacturing lots under 120 V peak overvoltage testing, 20% more than the datasheet maximum rating. All three devices were tested up to approximately 1.5 billion cycles, where a minimal $R_{DS(on)}$ shift was observed. The case temperature of all three DUTs was maintained at 75°C throughout the experiment by an active temperature control system. Due to the small junction-to-case thermal resistance of 0.5°C/W [72] and very little power dissipation during UIS testing (<0.3 W), the junction temperature of the DUT is virtually identical to the case temperature. As shown in Figure 5-6 (a), the in-situ measured $R_{DS(on)}$ in all cases is well below the datasheet limit scaled by the temperature coefficient (1.35x from 25°C to 75°C) [72]. In addition, the measured data points of each device follow a respective linear trend line in log-t scale on the horizontal axis, validating the lifetime model discussed in Section 4.2. Figure 5-6 (b) shows the 120 V overvoltage testing results of another representative 100 V rated GaN transistor EPC2302 [73] in a power quad flat no-lead (PQFN) package. The DUT was tested

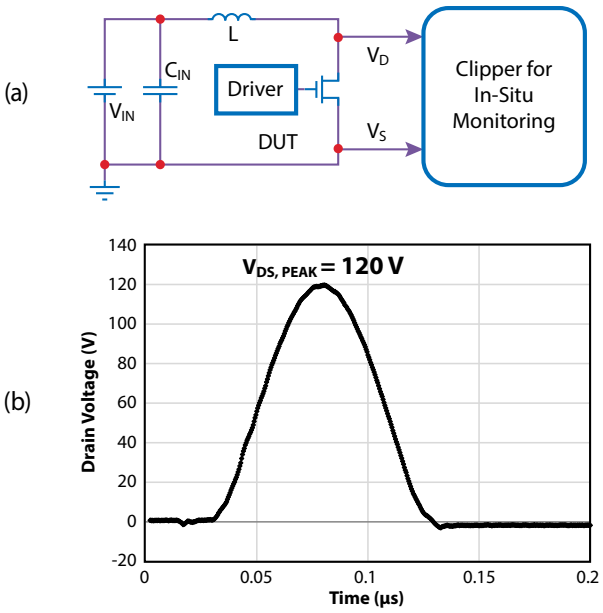


Figure 5-5 (a) Circuit schematic of a UIS test system with a clipper circuit used for in-situ $R_{DS(on)}$ monitoring. (b) a 120 V peak overvoltage drain waveform generated by UIS.

to approximately 10 billion cycles at ambient temperature (25°C), where very little $R_{DS(on)}$ shift was seen. A good agreement between 10 billion data points and the lifetime model (blue fit line) was also observed, proving the validity and versatility of the lifetime model. Results presented in Figure 5-6 show excellent overvoltage robustness of GaN HEMTs under 120% of $V_{DS,Max}$.

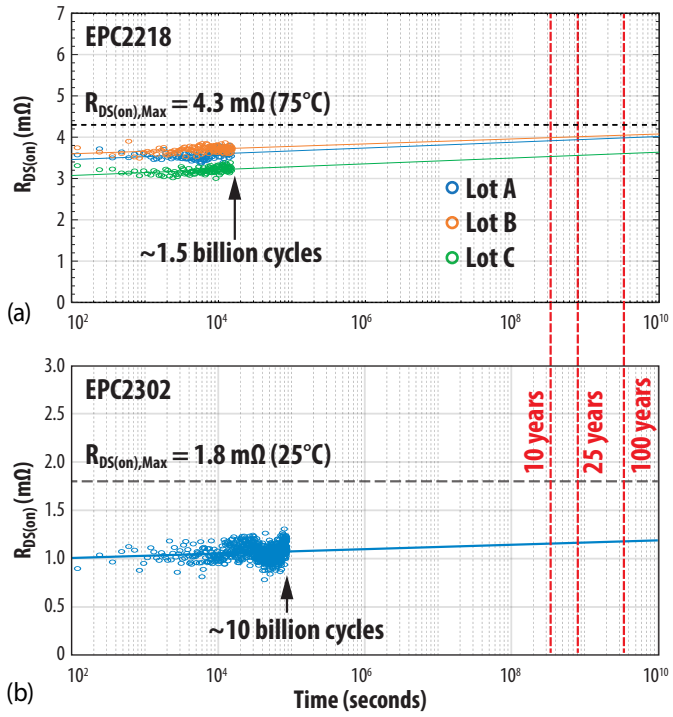
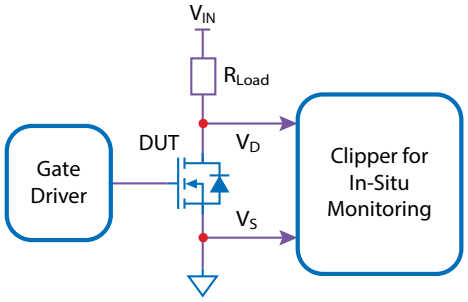


Figure 5-6 (a) In-situ measured $R_{DS(on)}$ from three different EPC2218 lots tested by UIS to 1.5 billion cycles of 120 V overvoltage spikes. (b) In-situ measured $R_{DS(on)}$ of a representative EPC2302 QFN GaN HEMT to 10 billion cycles of 120 V overvoltage spikes.

In a typical turn-off voltage waveform, there are usually multiple overvoltage oscillations before it stabilizes at the bus voltage. However, the first spike typically has the highest voltage. First-principles modeling estimates that the very first overvoltage pulse causes the most trapped charges, which dominates the dynamic $R_{DS(on)}$ shift in every switching period [1,27]. Therefore, the dynamic $R_{DS(on)}$ impact resulting from a single overvoltage pulse stress from UIS is representative of the entire ringing portion during a switching period.

Figure 5-4(c) shows how the equilibrium portion of the voltage waveform can be fitted. In Figure 5-7(a), a resistive hard switching topology circuit with in-situ $R_{DS(on)}$ monitoring was developed to study the wear-out mechanism involving hot electron trapping during hard switching. Figure 5-7(b) shows that the measured drain voltage rises from zero to the bus voltage (80 V) while the drain current (not presented) drops from the load current (several Amps) to virtually zero (leakage current) simultaneously. This hard-switched topology provides orders of magnitude more hot electrons than the typical high temperature reverse bias (HTRB) reliability testing configuration where the available number of electrons is limited by the low leakage current. The resistive load switching circuit also operates at 100 kHz with 15% duty cycle during which the DUT is on and $R_{DS(on)}$ is measured in-situ. This also means that the DUT is turned off 85% of the time, which is equivalent of 8.5 µs per switching period. Figure 5-7(b) plots the resulting hard switched turn-off voltage waveform that is matching the deconvoluted voltage waveform shown in Figure 5-4(c).

(a)



(b)

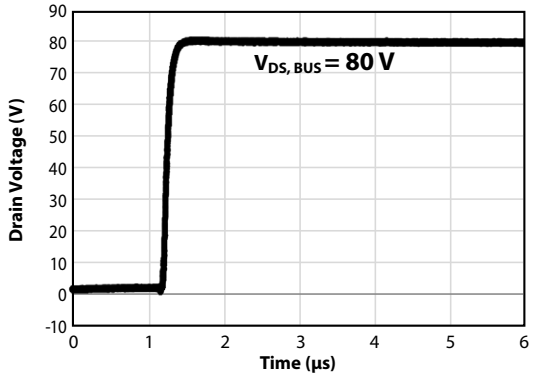


Figure 5-7 (a) Circuit schematic of a resistive load hard switching test system with a clipper circuit used for in-situ $R_{DS(on)}$ monitoring. (b) a turn-off drain voltage waveform to 80 V bus voltage produced by the resistive load hard switching circuit.

Figure 5-8 shows the test results of one representative of EPC2218 and EPC2302 each under 80 V, 100 kHz testing condition. To better view the evolution of $R_{DS(on)}$ drift, all the in-situ measured $R_{DS(on)}$ were normalized to the first measured data point and plotted in Figure 5-8 where the vertical axis is normalized $R_{DS(on)}$. Similar to the UIS results, the lifetime model also provides a good fit to the data points collected by the resistive load hard switching test circuit, which further validates the applicability of the lifetime model. The model predicts less than 10% $R_{DS(on)}$ increase over 100 years of continuous switching at 100 kHz and 80 $V_{DS,Bus}$, as shown in Figure 5-8, revealing good robustness of GaN HEMTs under nominal bus voltage hard-switched stress conditions.

Previous work also conclusively demonstrated that this hot electron trapping induced $R_{DS(on)}$ shift has a negative temperature coefficient because of the negative temperature dependence of mean free path discussed in Section 4.2.

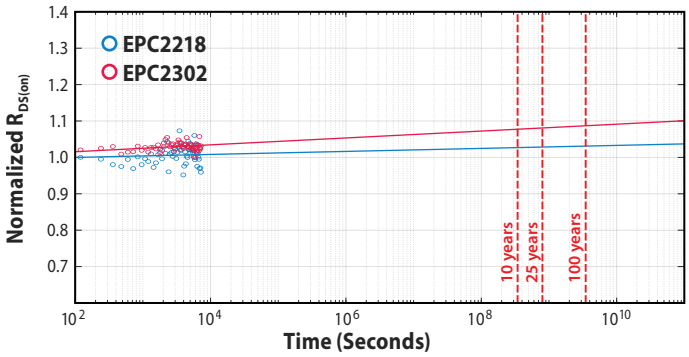


Figure 5-8: The in-situ measured $R_{DS(on)}$ of one EPC2218 and one EPC2302 under 80 V and 100 kHz resistive load had-switched testing conditions, where both devices project less than 10% $R_{DS(on)}$ shift over 25 years of continuous operation.

Using this information, how can the reliability results from two different testing topologies be combined into one that is representative of a real-world DC-DC converter?

Because two different testing topologies address different spectrums of a common turn-off voltage waveform from a buck converter, the reliability impact of each individual stressor can be combined as shown in Equation 5-6, which highlights that the harsher drain bias stressor dominates the overall lifetime.

$$\frac{1}{MTTF_{Total\_Drain}} = \frac{1}{MTTF_{Overvoltage}} + \frac{1}{MTTF_{Bus\ Voltage}} \qquad \text{Eq. 5-6}$$

Previously, 25 years of continuous operation was used as a lifetime projection target used in Figures 5-6 and 5-8 for general DC-DC converter applications. However, the projected $R_{DS(on)}$ values at the end of 25 years are still notably less than the datasheet maximum limit in both cases.

Therefore, a more stringent failure criterion is implemented here to determine time-of-failure for each and combined projected lifetime results. A 20% in-situ $R_{DS(on)}$ drift compared to the first read point is used to estimate the time-of-failure for EPC2218 for respective testing circuits. Figure 5-9 shows the projected time-to-failure for

EPC2218 under UIS (120 V $V_{DS,Peak}$) and resistive load hard switching (80 $V_{DS,Bus}$) is 8 x 10^{10} seconds and 4 x 10^{15} seconds, respectively. By plugging the time-of-failure results into Equation 5-4, the total lifetime is dominated by the overvoltage contribution because it is orders of magnitude less than the resistive load switching testing result. The total lifetime is calculated to be approximately 2,570 years, which is based on 100 kHz testing data. If designers need to scale the projected results to the actual operating frequency, a simple frequency ratio can be applied to adjust the lifetime as discussed earlier, where 1 MHz operating frequency would yield 257 years of equivalent lifetime.

The projected total lifetime results show that even under an extreme drain bias condition caused by a buck converter with severe overshoot, GaN HEMTs still demonstrated excellent robustness. In summary, dynamic on-resistance wear out mechanism should not be a critical concern for EPC's GaN HEMTs for use in common DC-DC converters.

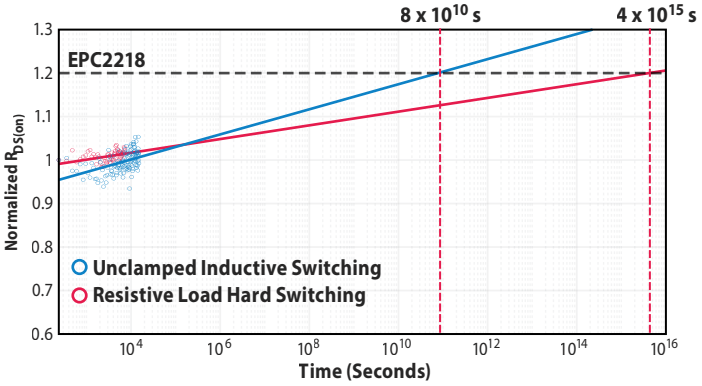


Figure 5-9 Normalized $R_{DS(on)}$ of two EPC2218 devices were projected to a time where $R_{DS(on)}$ shifts 20% as compared to first read point. One EPC2218 device was tested by UIS test circuit. The other one was tested by the resistive load hard switching circuit.

5.2.5. Temperature Cycling

Temperature cycling is another critical area of interest for DC-DC converter applications.

This analysis is based on the board-level thermomechanical reliability study presented in Section 4.4.4, which showed that proper underfill material improves the temperature cycling lifetime of CSP GaN devices by a factor of at least 4.8x. In the following discussions, only TC1 with underfill data is used.

For an upper limit in this analysis, T_{Max} is assumed to be 125°C, which is the typical maximum design temperature for power modules. The number of cycles to failure (N) at 100 ppm, or 0.01%, failure rate for EPC2218A with underfill can be plotted as a function of ΔT using Equation 4-19 (Section 4.4.4), while the Arrhenius term is a constant coefficient. The result is shown by the black line in Figure 5-10. The horizontal axis (ΔT) only includes a range of 0 to 100°C because power modules in real-world applications are typically kept at 25°C ambient temperature when not in operation, which yields a maximum ΔT of 100°C.

In some of the DC-DC converters that are designed for a lower T_{Max} of 100°C during normal operation, the Arrhenius term should now be slightly larger due to a smaller denominator (T_{Max}) in the exponential equation. The red line in Figure 5-10 shows the number of cycles to fail at 100 ppm extracted from the Weibull distribution as a function of ΔT, where the red curve is slightly above the black curve (T_{Max} = 125°C). Because T_{Max} is lowered by 25°C, the red curve is now plotted from 0°C to 75°C on the horizontal ΔT axis.

For some applications that are designed for a T_{Max} of 75°C, the model is plotted in blue, where a longer lifetime is expected because of the larger Arrhenius term. A T_{Max} of 50°C is also included in Figure 1, as shown in the yellow line.

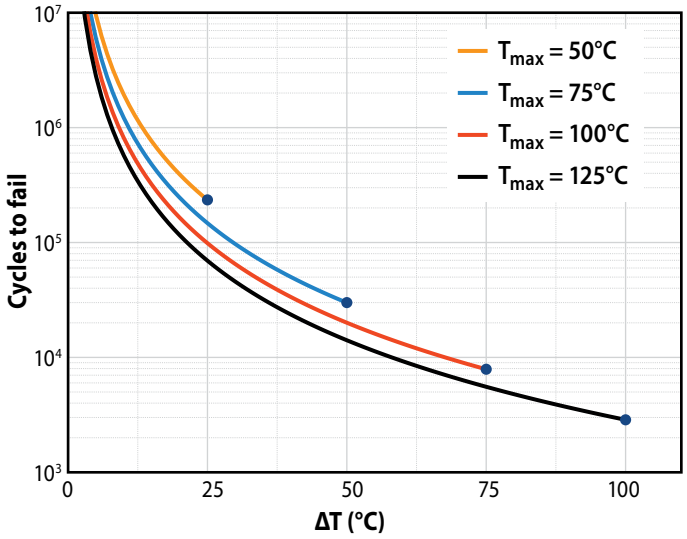


Figure 5-10. Number of cycles to fail at 100 ppm or 0.01% failure rate vs. ΔT at T_{Max} of 50°C (yellow), 75°C (blue), 100°C (red), and 125°C (black).

How can designers use Figure 5-10 to determine the TC lifetime for their DC-DC converter design?

By way of example, take a converter that will be operating in the desert climate of Phoenix, AZ, USA. The ambient outside temperature in the summer can be as high as 50°C (122°F). This notional converter generates another 75°C of heat during operation, which gives a T_{Max} of 125°C. By following the black curve in Figure 5-10 and finding the vertical intercept where ΔT of the horizontal axis is 75°C, the estimated number of cycles to 100 ppm failure rate is a little more than 5000 cycles, hopefully representing decades of operation when also considering the more moderate temperature seasons. This approach provides a practical method to correlate lab generated TC reliability results to real-world applications.

5.2.6. Conclusions

After reviewing the common stresses experienced by DC-DC converters, a test-to-fail approach was adopted and applied to investigate the intrinsic underlying wear-out mechanisms of GaN HEMTs. Three stressors that are most likely responsible for

device failures are identified, which are gate bias, drain bias and temperature cycling. Under gate bias, a physics-based lifetime model based on impact ionization was used to predict the lifetimes. A 1 ppm failure rate was projected after 25 years of continuous DC gate bias at the maximum rated voltage (V_{GS} = 6 V). Another physics-based model based on hot electron trapping mechanism was used to explain the dynamic $R_{DS(on)}$ wear-out mechanism under drain bias. The measured data and the lifetime model predict that the $R_{DS(on)}$ shift is expected to be less than 20% over the lifetime of the part. The wear-out mechanism responsible for temperature cycling (TC) failure is solder joint cracking. A third lifetime model that includes TC range, temperature extreme, and cycling speed was introduced. Combining the wear-out rates of all three stressors shows that neither gate bias nor drain bias is of significant reliability concern in DC-DC converter applications. Thermo-mechanical stress due to TC is deemed to have the highest risk that warrants careful considerations. Using appropriate underfill materials can vastly reduce TC reliability risk while giving excellent lifetimes.

5.3. Lidar Application Reliability

5.3.1. Introduction to Lidar Reliability

Compared to other applications, GaN FETs used for light detection & ranging (lidar) are often subject to long durations of reverse bias and short pulses of relatively high current. This section evaluates the reliability of devices used in lidar applications, both discrete FETs and GaN lidar ICs which include low-voltage driver circuits.

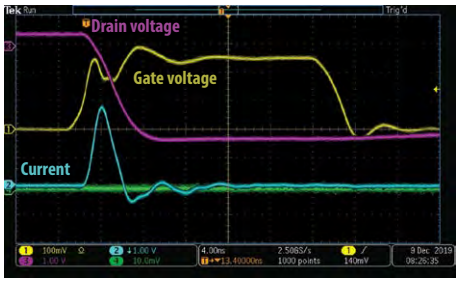
5.3.2. Long-Term Stability Under High Current Pulses

The concept of this test method is to stress parts in an actual lidar circuit for a total number of pulses well beyond their ultimate mission profile. The mission profiles for automotive lidar vary from customer to customer. A typical automotive profile would call for a 15-year life, with two hours of operation per day, at 100 kHz pulse repetition frequency (PRF). This corresponds to approximately four trillion total lidar pulses. Some worst-case (heavy use) scenarios might call for as many as 10–12 trillion pulses in service life.

By testing a population of devices well beyond the end of their full mission profile while verifying the stability of the system performance and the device characteristics, this test method directly establishes the suitability of eGaN devices for lidar applications. To achieve the large number of pulses, parts are stressed continuously, rather than in bursts as used in typical lidar circuits.

For this study, two popular AEC grade parts were put under test: EPC2202 (80 V) and EPC2212 (100 V). Four parts of each type were tested simultaneously. During the stress, two key parameters were continuously monitored on every device: (1) peak pulse current and (2) pulse width. These parameters are both critical to the range and resolution of a lidar system.

Figures 5-11 and 5-12 show the results of this test over the first 13 trillion pulses. The cumulative number of pulses well exceeds a typical automotive lifetime and covers worst-case use conditions. Note that there is no observed degradation or drift in either the pulse width or height. While this is an indirect monitor of the health of the GaN device, it indicates that no degradation mechanisms have occurred that would adversely impact lidar performance.



AEC-Q101 series of discrete FETs

- 8 samples (>7000h)
- 0 failures and perfect pulse stability

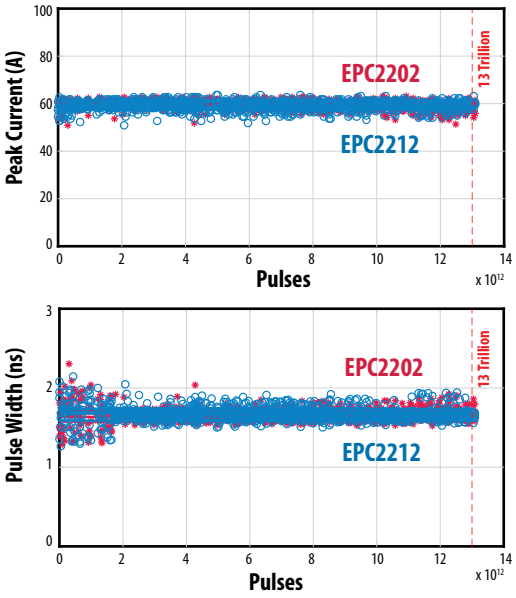


Figure 5-11: Long-term stability of pulse width (bottom right) and pulse height (top right) over 13-trillion lidar pulses. Data for four EPC2202 (red) devices and four EPC2212 (blue) devices are overlaid in the plots.

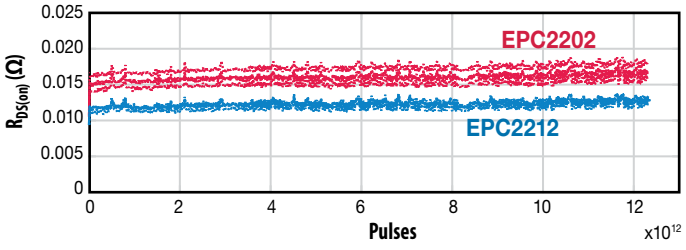


Figure 5-12: Long-term stability of $R_{DS(on)}$ during lidar reliability testing. These parameters are measured at six-hour intervals on every part by briefly interrupting the lidar stress. Data for four EPC2202 (red) devices and four EPC2212 (blue) devices are overlaid in the plots.

5.3.3. Monolithic GaN-on-Si Laser Driver ICs

Lidar systems often use discrete eGaN transistors separate from a gate driver chip due to the benefits of GaN's small footprint and superior switching performance. EPC recently introduced a family of GaN laser drive IC products that integrate a high-speed GaN driver with the discrete GaN transistor (see Figure 5-13). This integrated monolithic lidar solution offers even higher performance, smaller form factor, and lower cost than the existing discrete solutions. As a result, these ICs enable a wider range of lidar applications including robotics, surveillance systems, drones, autonomous cars, vacuum cleaners, and many more.

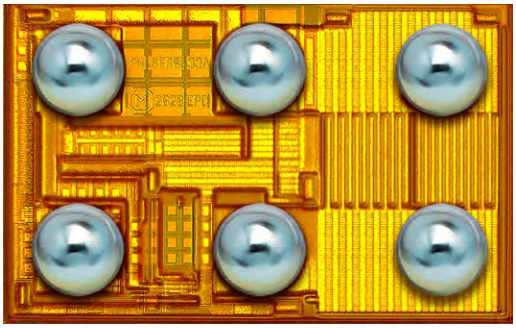


Figure 5-13: The EPC21601 eToF™ integrated circuit includes a driver and a power FET.

The first two offerings of the integrated GaN laser drive IC products (EPC21601 and EPC21701) are in production. Table 5-1 summarizes the main specifications of the first two qualified IC products.

Part Number	Die Size (mm x mm)	Main Specifications
EPC21601	S (1.5 X 1)	40 V, 15 A, 3.3 V logic, eToF laser driver IC
EPC21701	S (1.7 X 1)	80 V, 15 A, 3.3 V logic, eToF laser driver IC

Table 5-1: Initial EPC Laser Driver IC Product Family

5.3.4. Key Stressors of eToF Laser Driver IC for Lidar Application

The integration of the gate driver and power transistor into a chip-scale package greatly reduces the parasitic inductances and further improves the speed, minimum pulse width and power dissipation. It also introduces challenges in isolating the key electrical stressors because many of the IC's voltages and currents cannot be accessed directly. The first step of the study is to identify the key stressors that affect the IC in lidar applications.

Both EPC21601 and EPC21701 are selling in a chip-scale BGA form factor that measure at 1.5 mm x 1.0mm and 1.7 mm x 1.0 mm, respectively. The package technology of the laser driver ICs has been used in EPC's discrete power transistors for many years, and therefore the package related reliability of the IC products was covered by previous phase reliability testing reports and related publications [1, 15, 27, 78, 79, 80].he lidar IC's operating conditions, shown in Figure 5-14, are best emulated through High Temperature Operating Life (HTOL) testing. EPC21601 is selected as the test vehicle for this test-to-fail study as it was released a few months earlier than EPC21701. The laser driver circuit design of the two products is identical. The main difference between them is the drain voltage rating of the output GaN transistor, where EPC21601 has an absolute V_D max rating of 40 V and EPC2701 is 80 V.

Three key stressors are identified:

- Logic supply voltage V_{DD} that supplies the drive voltage to the low voltage (LV) GaN FETs in laser driver circuit as well as the gate of the high voltage (HV) GaN output FET.
- Laser drive voltage V_D that is predominantly applied to the drain terminal of the HV output FET.
- Operating frequency which stresses both the LV laser driver circuits and the HV output FET.

5.3.5. Effect of V_{DD}, Logic Supply Voltage

When EPC21601 is operated and generates a burst of short pulses, the logic supply voltage (V_{DD}) is applied to the gate terminals of the LV GaN FETs in the laser driver circuits and the gate of the HV GaN power transistor. It is equivalent of performing a dynamic gate test for all GaN FETs with a burst frequency of 1 kHz, very low duty cycle (~0.02%), and high operating frequency (30 MHz). When not pulsed, the part is in the OFF state and the gate bias is nearly zero (see Figures 5-14 and 5-15).

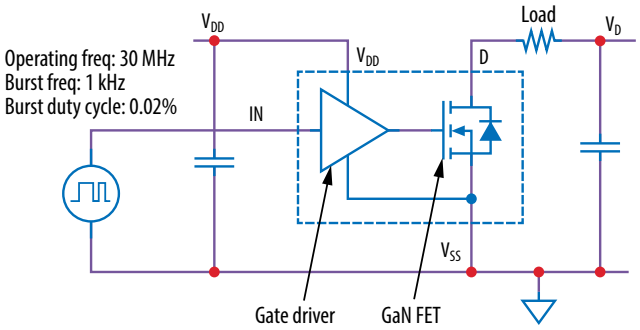


Figure 5-14: Block diagram of EPC21601 and EPC21701 laser drive integrated circuits.

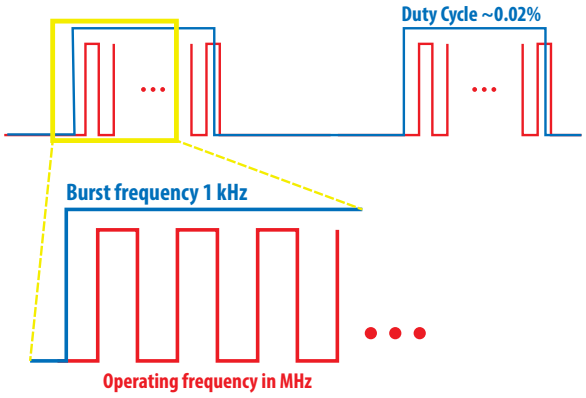


Figure 5-15: Diagram of operating conditions with Burst Frequency (Blue) 1 kHz with a duty cycle of ~0.02% and Operating Frequency in MHz

In the qualification HTOL test, V_{DD} was biased at the absolute maximum rating of 5.5 V, and no issue was found after 1000 hours of testing at 125°C junction temperature. To test the device's robustness, the V_{DD} voltage was increased to a high value at 7 V, which is more than 125% of the absolute maximum rating. This stress condition addresses the worst overvoltage ringing issue on the V_{DD} pin during normal operation by customers. Table 5-2 summarizes the test result where 16 devices were tested up to 1049 hours at 7 V V_{DD} and 125°C junction temperature. No failures occurred. This indicates that a significant margin exists in the laser driver IC products.

Stress Test	Part Number	Test Condition	# of Failure	Sample Size	Duration (Hrs)
HTOL	**EPC21601**	**V_{DD} = 7 V, T_J = 125°C**, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	0	16	1049

Table 5-2: HTOL Test Result of EPC21601 with V_{DD} = 7 V and T_J = 125°C

As there were zero failures, this result does not determine how much margin was designed into the product or to accurately predict the lifetime at a given mission profile for the V_{DD} stressor. Therefore, more stringent stress conditions must be applied to test the devices to failure, where the goal is to fail the parts quickly and conduct failure analysis to understand the underlying failure modes and mechanisms.

To determine the voltage acceleration of the V_{DD} stress, a matrix of tests was conducted from 8.5 V to 9.5 V at 25°C, as shown in Table 5-3. At 8.5 V V_{DD}, a total of three failures were found after more than 1000 hours of testing whereas almost all parts failed within 305 hours at 9.5 V, indicating a significant voltage acceleration.

Stress Test	Part Number	Test Condition	# of Failure	Sample Size	Duration (Hrs)
HTOL	**EPC21601**	**V_{DD} = 8.5 V, T_J = 25°C**, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	3	16	1049
HTOL	**EPC21601**	**V_{DD} = 9.5 V, T_J = 25°C**, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	15	16	305

Table 5-3: HTOL Test Result of EPC21601 with V_{DD} = 8.5 V and V_{DD} = 9.5 V, T_J = 25°C

Temperature acceleration was also studied by conducting HTOL tests at 25°C and 125°C, while the V_{DD} was fixed at 8.5 V. The results are summarized in Table 5-4 where it shows a significant temperature acceleration.

Stress Test	Part Number	Test Condition	# of Failure	Sample Size	Duration (Hrs)
HTOL	**EPC21601**	**V_{DD} = 8.5 V, T_J = 25°C**, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	3	16	1049
HTOL	**EPC21601**	**V_{DD} = 8.5 V, T_J = 125°C**, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	16	16	718

Table 5-4: HTOL Test Result of EPC21601 with T_J = 25°C and T_J = 125°C, V_{DD} = 8.5 V

Failure analysis determined that all failures were soft parameter failures in which quiescent current exceeded the 20 mA maximum datasheet limit, with V_{DD} = 5 V and the measurement conducted during the OFF state [81]. Under closer examination, the quiescent current only exceeded datasheet limits when V_D = 20 V was provided. When the quiescent current soft failures were subjected to lidar operation with a V_D of 15 V, the integrity of their pulses was uncompromised. Figure 5-16 shows the waveforms of the input signal (blue) of V_{IN} (the logic input to EC21601) and the corresponding output signals from V_D of the quiescent current failures (green and yellow), where no pulse distortion or missing pulses were observed. This suggests even when the device was damaged by extremely high V_{DD} stress, it still was functional, and the repeatability of current pulses was not adversely impacted.

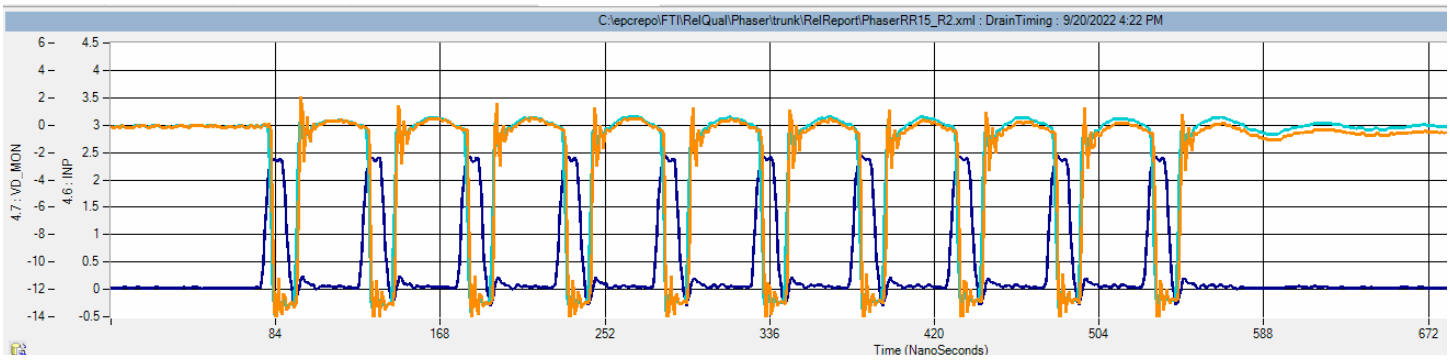


Figure 5-16: The input (blue) waveform and the corresponding output waveforms of the quiescent current failures (green and yellow)

Since all failures at different voltages and temperatures showed similar "soft" electrical failures, physical failure analysis was conducted to determine the underlying root cause. Gate rupture of the LV GaN FETs in the driver circuit was found to be the single failure mechanism for all failures regardless of stress voltages and temperatures. This result is expected based on the circuit analysis because the V_{DD} voltage is applied to the gates of the LV and HV GaN FETs when the pulses are generated.

Figure 5-17 shows time-to-failure data for the two different V_{DD} voltages at room temperature. The data was analyzed using a two-parameter Weibull distribution for each voltage leg using maximum likelihood estimation (MLE). The fits are indicated by solid lines in the graphs. The Weibull shape (or slope) parameter was constrained to be the same for all voltage legs because a single failure mode was found through failure analysis.

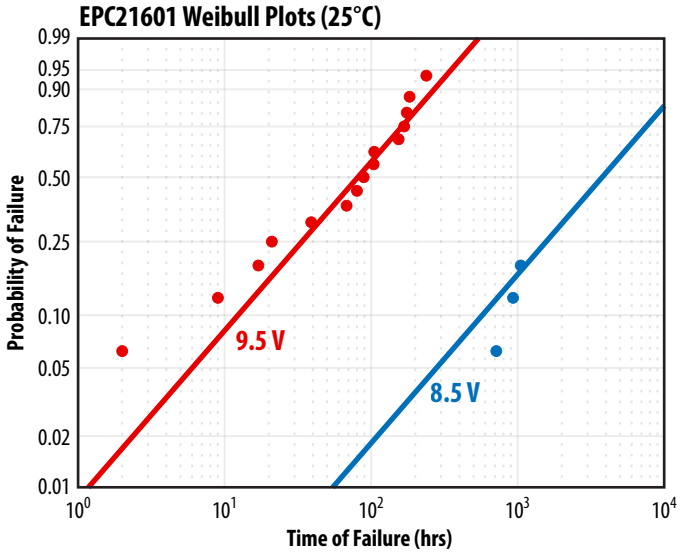


Figure 5-17: Weibull plots showing the failures of EPC21601 at 8.5 V (blue) and 9.5 V (red) V_{DD}, respectively and T_J = 25°C.

The calculated mean-time-to-failure (MTTF) of the 9.5 V V_{DD} leg is approximately 117 hours, which equals 4.2 x 10^5 seconds. In Figures 1 and 2 of the Phase 14 Reliability Report [60], the MTTF of the 9.5 V V_{GS} DC test of EPC2212 at 25°C is approximately 150 seconds, which is 7.5 x 10^5 seconds when scaling with the 0.02% burst duty cycle that was used in the HTOL test. EPC21601 and EPC2212 share the same gate construction and use identical gate fabrication processes. This shows that static DC V_{GS} testing on EPC2212 and the measured MTTF of EPC21601 in accelerated dynamic gate testing are consistent. It is understandable that the two MTTF values do not match exactly due to the difference in testing setup and implementation. For instance, the gates of all the LV FETs were stressed through the same V_{DD} pin concurrently during an extremely short pulse, where some slight ringing on the gates might be expected. This could explain the slightly worse MTTF for EPC21601 as compared to the DC accelerated gate testing result for EPC2212.

The commensurate MTTF results between EPC21601 and EPC2212 also corroborate the validity of the physics-based model EPC developed for the gate reliability. The same lifetime model fits the measured data for V_{DD} at both biases.

Figure 5-18 shows the lifetime projection against the measured acceleration data for EPC21601 at 25°C. The fit projected greater than 25 years of lifetime with less than 1 ppm failure rate at the 5.5 V maximum V_{DD} voltage rating at 25°C. This result also agrees well with the extrapolated lifetime for gate at 5.5 V under static DC gate bias.

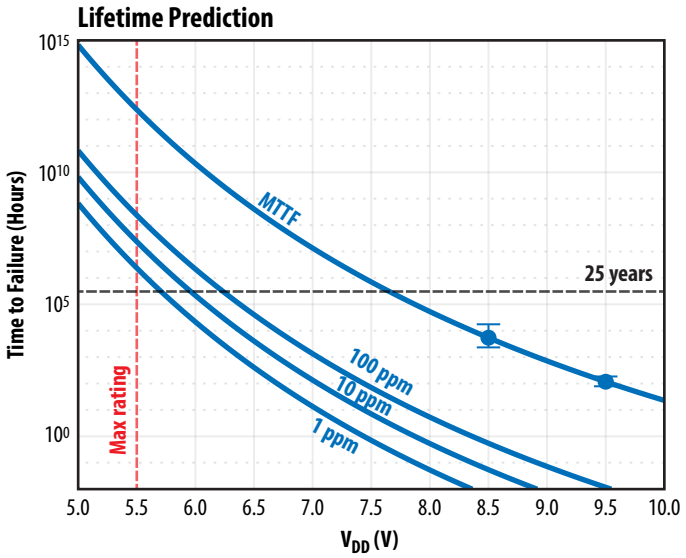


Figure 5-18: EPC21601 MTTF data at two different voltages with error bars are plotted against V_{DD} at 25 °C. The solid line corresponds to the impact ionization lifetime model. Extrapolations of time to failure for 100 ppm, 10 ppm, and 1 ppm are shown as well.

Temperature acceleration of the time-to-failure data are shown in Figure 5-19 (25°C and 125°C) while V_{DD} was fixed at 8.5 V. The data was also analyzed using a two-parameter Weibull distribution for each temperature leg using maximum likelihood estimation (MLE). The Weibull shape (or slope) parameter was constrained to be the same for both temperature legs because a single failure mode was identified through failure analysis. The time-to-fail of each device was recorded by conducting a complete ATE post screening after the parts were removed from the oven (125°C leg) and the motherboards. Multiple "soft" quiescent current failures were found at the same first read point at 72 hours in the 125°C leg, where a cluster of vertical failure data points were shown on the Weibull plot. The last failure was found at 718 hours for the 125°C leg, whereas only a total of three soft failures were measured after more than 1000 hours of testing in the 25°C leg, as shown in Table 5-4.

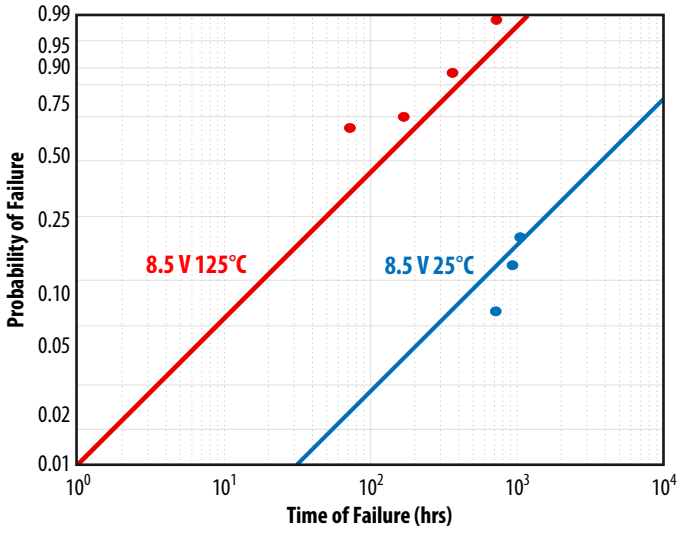


Figure 5-19: Weibull plots showing the failures of EPC21601 at 25°C (blue) and 125°C (red) junction temperature, V_{DD} = 8.5 V.

Figure 5-20 shows the Arrhenius plot for the MTTF data at 25°C and 125°C with V_{DD} = 8.5 V, where an activation energy of 0.35 eV was calculated by using the Arrhenius equation [82, 83, 84]. This result is different from what was observed when conducting static HTGB testing for discrete GaN products, which showed weak negative temperature acceleration. Initial failure analysis showed identical gate rupture as the underlying failure mode for all soft quiescent current failures regardless of 25°C or 125°C testing temperature.

Though the failure mechanism responsible for the temperature acceleration warrants further investigation, the laser driver IC under the V_{DD} stressor is proven to be extraordinarily robust.

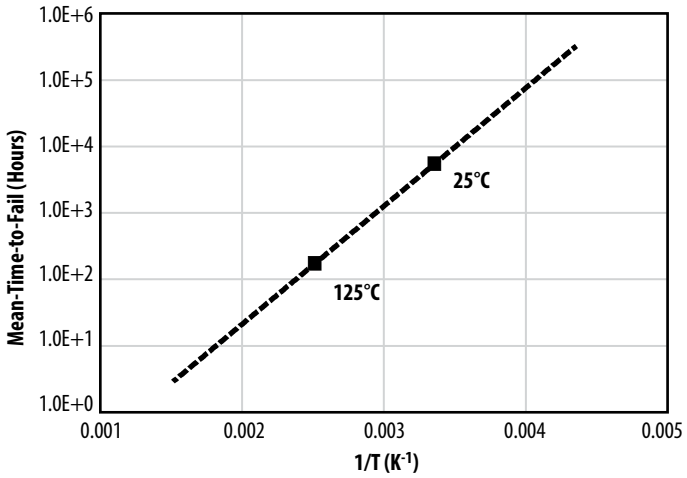


Figure 5-20: EPC21601 MTTF data at two different temperatures are plotted against T^{-1} (K^{-1}) with V_{DD} at 8.5 V. The solid line corresponds to the Arrhenius equation, where an activation energy of 0.35 eV was found.

5.3.6. Effect of V_D, Laser Drive Voltage

By examining the circuits that connect to the V_D pin in detail, the accelerated V_D HTOL can cause two potential failure modes in EPC21601.

1. V_D primarily goes to the drain terminal of the HV GaN FET. Due to the nature of lidar operation, the HV output FET is under reverse drain bias most of the time. When the laser pulses are generated, the HV FET turns on and conducts current. Accelerated V_D HTOL testing of the IC therefore resembles a dynamic HTRB test of the output FET with a high duty cycle. Therefore, the intrinsic failure modes due to accelerated drain bias test for a discrete GaN transistor apply.
2. Besides connecting to the drain node of the HV FET, the V_D pin also connects to a single laser driver circuit, which affects the number of pulses generated by the device. If that path was compromised by the accelerated V_D stress, it could lead to missing pulses, which is another crucial failure mode for lidar application.

The HTOL qualification test was conducted at 30 V V_D, the maximum recommended voltage specified by the datasheet [85]. A matrix of accelerated V_D HTOL tests were conducted as summarized in Table 5-5. 60 V V_D was selected because it is two times of the maximum recommended voltage rating, which is an extremely accelerated condition. However, this voltage is not too high to cause some other known intrinsic failure modes for the HV output FET. 60 V is an aggressive test-to-fail condition against the driver design.

Table 5-5 shows that no failures were found after more than 1000 hours of testing. All parts continued to meet the datasheet specifications after undergoing the HTOL tests.

Stress Test	Part Number	Test Condition	# of Failure	Sample Size	Duration (Hrs)
HTOL	EPC21601	V_{DD} = 5.5 V, **T_J = 25°C** **V_{D_DC} = 60 V** V_{IN} = 3.3V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	0	16	1005
HTOL	EPC21601	V_{DD} = 5.5 V, **T_J = 125°C** **V_{D_DC} = 60 V** V_{IN} = 3.3V_{P-P}, Burst frequency = 1 kHz; Operating frequency = 30 MHz	0	16	1005

Table 5-5: HTOL Test Result of EPC21601 with V_D = 60 V, T_J = 25°C and T_J = 125°C, respectively

To further validate that the devices were not generating distorted waveforms or missing pulses, the parts from the V_D = 60 V and T_J = 125°C leg were mounted back onto the test setup at 60 V and 125°C and the input and output pulse waveforms were captured as shown in Figure 5-21. This result shows that no degradation in pulse waveforms was observed after more than 1000 hours of HTOL testing. It is also important to note that the HV output transistor experienced more than 25 V overshoot at the end of each pulse during HTOL resulting from the short pulses. It suggests that the device saw repetitive > 85 V transient overvoltage stress (> two times the absolute maximum rating = 40 V) on V_D in addition to the 60 V nominal stress that is another two times the maximum recommended bias. This also demonstrates good robustness of the device under V_D stress.

At this point, the most rigorous testing corner is covered by the testing matrix at the 60 V V_D leg at 125°C. Further increasing the drain bias might introduce a different intrinsic failure mechanism for the HV GaN transistor that is not applicable to the lidar application or the reliability of the laser drive IC. In short, no failure mode was found to be associated with the laser supply voltage (V_D).

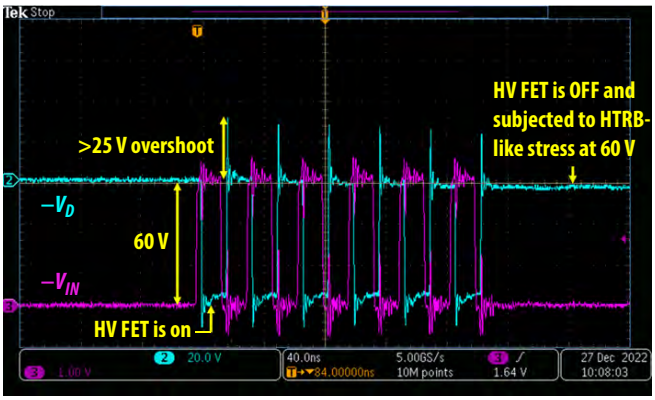


Figure 5-21: Output waveforms (blue) of a representative passing part after it was subjected to 1005 hours of HTOL testing at 60 V V_D and 125°C. The purple waveform is the corresponding input signal from V_{IN}. Please note that a 25 V of overshoot was seen at the end of each pulse during HTOL testing.

5.3.7. Effect of Operating Frequency

Preliminary device characterization suggested that the output waveforms of lidar ICs could be distorted when tested at extremely high operating frequencies. It is therefore useful to study at what frequency or duration of the HTOL testing the pulse waveform starts showing significant distortion or missing pulses.

Tests at two high operating frequencies were carried out as shown in Table 5-6. 48 MHz and 96 MHz are 160% and 320% of the 30 MHz maximum recommended operating frequency used in qualification. No failure occurred after more than 1400 hours of testing. All parts continued to meet the datasheet specifications after undergoing the HTOL tests.

Figure 5-22 shows representative input (purple) and output (blue) waveforms of a passing device post 1413 hours of 48 MHz HTOL testing. No waveform distortion or missing pulses were found. Figure 5-23 shows another set of representative input (purple) and output (blue) waveforms of a passing device post 1413 hours of 96 MHz HTOL testing. No waveform distortion or missing pulses were found.

Stress Test	Part Number	Test Condition	# of Failure	Sample Size	Duration (Hrs)
HTOL	EPC21601	V_{DD} = 5.5 V, T_J = 25°C, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; **Operating frequency = 48 MHz**	0	16	1005
HTOL	EPC21601	V_{DD} = 5.5 V, T_J = 25°C, V_{D_DC} = 30 V, R_{LOAD} = 2 Ω V_{IN} = 3.3 V_{P-P}, Burst frequency = 1 kHz; **Operating frequency = 96 MHz**	0	16	1005

Table 5-6: HTOL Test Result of EPC21601 with operating frequency of 48 MHz and 96 MHz with V_D = 30 V and T_J = 25°C.

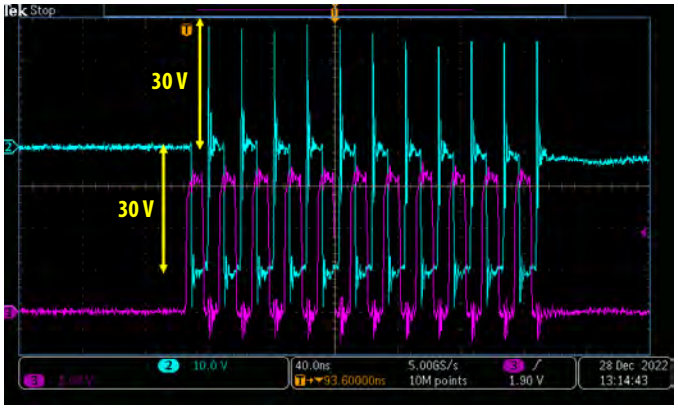


Figure 5-22: Representative input (purple) and output (blue) waveforms of a passing device after 1413 hours of HTOL testing at 48 MHz operating frequency. Please note that a 30 V of overshoot was seen at the end of each pulse during HTOL testing. The device continues to pulse without distortion at 48 MHz.

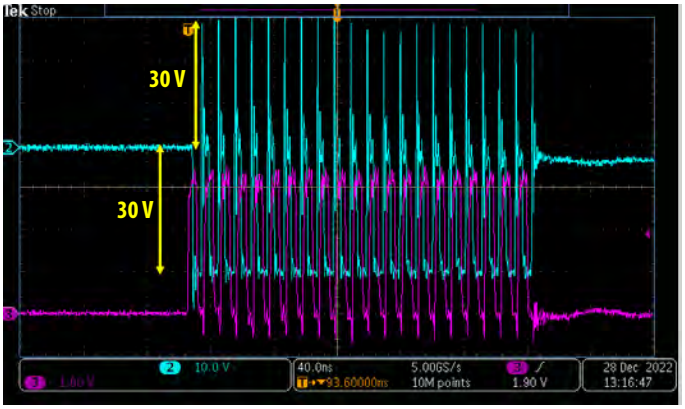


Figure 5-23: Representative input (purple) and output (blue) waveforms of a passing device after 1413 hours of HTOL testing at 96 MHz operating frequency. Please note that a 30 V of overshoot was seen at the end of each pulse during HTOL testing. The device continues to pulse without distortion at 96 MHz.

6.0. SUMMARY AND CONCLUSIONS

As GaN device production continues to increase and applications diversify, separate reliability concerns arise which may depend on the use case. By understanding the wear-out mechanisms that affect a system in each phase of its mission profile, GaN device lifetimes can be calculated analytically for each specific application. The failure rate of each wear-out mechanism, which is confirmed by testing to failure, can be minimized by following the guidelines provided in this report.

7.0. REFERENCES

[1] Garcia, R., Gajare, S., Espinoza, A., Zafrani, M., Pozo, A., Zhang, S., "GaN Reliability and Lifetime Projections: Phase 15", EPC Corp., El Segundo, CA, USA, Reliability Report. Available: Reliability Report Phase 15 (epc-co.com)

[2] Handbook for Robustness Validation of Semiconductor Devices in Automotive Applications, Third edition: May 2015, Editor: ZVEI Robustness Validation Working Group, Eds. Published by ZVEI – Zentralverband Elektrotechnik – und Elektronikindustrie e.V. [Online]. Available: https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Publikationen/2015/mai/Handbook_for_Robustness_Validation_of_Semi- conductor_Devices_in_Automotive_Applications__3rd_edition_/Robustness-Validation-Semiconductor-2015.pdf

[3] Volosencu, C. (2017). System reliability. Edited. InTechOpen. https://doi.org/10.5772/66993.

[4] JEDEC Standard "Method for calculating failure rates in units of FITS", JESD85A, January 2014.

[5] Gao, H., Espinoza, A., Garcia, R., Gajare, S., & Zhang, S. (2024). Main Mechanism Responsible for pGaN Gate Breakdown and Lifetime Projection in GaN HEMTs. IEEE Transactions on Electron Devices.

[6] Rossetto, I., Meneghini, M., Hilt, O., Bahat-Treidel, E., De Santi, C., Dalcanale, S., ... & Meneghesso, G. (2016). Time-dependent failure of GaN-on-Si power HEMTs with p-GaN gate. IEEE Transactions on Electron Devices, 63(6), 2334-2339.

[7] Cao, L., Wang, J., Harden, G., Ye, H., Stillwell, R., Hoffman, A. J., & Fay, P. (2018). Experimental characterization of impact ionization coefficients for electrons and holes in GaN grown on bulk GaN substrates. Applied Physics Letters, 112(26).

[8] Okuto, Y., & Crowell, C. R. (1975). Threshold energy effect on avalanche breakdown voltage in semiconductor junctions. Solid-State Electronics, 18(2), 161-168.

[9] Cao, L., Zhu, Z., Harden, G., Ye, H., Wang, J., Hoffman, A., & Fay, P. J. (2021). Temperature dependence of electron and hole impact ionization coefficients in GaN. IEEE Transactions on Electron Devices, 68(3), 1228-1234.

[10] Gajare, S., Gao, H., Wong, C., & Zhang, S. (2024). A Comprehensive Lifetime Model for Schottky-type pGaN Gate of GaN HEMTs. IEEE Electron Device Letters.

[11] Richardson, O. W. (1924). Electron emission from metals as a function of temperature. Physical Review, 23(2), 153.

[12] He, J., Wei, J., Li, Y., Zheng, Z., Yang, S., Huang, B., & Chen, K. J. (2020). Characterization and analysis of low-temperature time-to-failure behavior in forward-biased Schottky-type p-GaN gate HEMTs. Applied Physics Letters, 116(22).

[13] Chiu, F. C. (2014). A review on conduction mechanisms in dielectric films. Advances in Materials Science and Engineering, 2014(1), 578168.

[14] Zhang, H., Miller, E. J., & Yu, E. T. (2006). Analysis of leakage current mechanisms in Schottky contacts to GaN and Al0. 25Ga0. 75N/GaN grown by molecular-beam epitaxy. Journal of Applied Physics, 99(2).

[15] Wang, Bixuan, Ruizhe Zhang, Hengyu Wang, Quanbo He, Qihao Song, Qiang Li, Florin Udrea, and Yuhao Zhang. "Gate Lifetime of P-Gate GaN HEMT in Inductive Power Switching." In 2023 35th International Symposium on Power Semiconductor Devices and ICs (ISPSD), pp. 20-23. IEEE, 2023

[16] Alex Lidow. "GaN Power Devices and Applications.", El Segundo, CA, USA: PCP Press, 2021

[17] Dymond, Harry CP, Jianjing Wang, Dawei Liu, Jeremy JO Dalton, Neville McNeill, Dinesh Pamunuwa, Simon J. Hollis, and Bernard H. Stark. "A 6.7-GHz active gate driver for GaN FETs to combat overshoot, ringing, and EMI." IEEE Transactions on Power Electronics 33, no. 1 (2017): 581-594

[18] Zhang, S., Gajare, S., Garcia, R., Huang, S., Espinoza, A., Gorgerino, A., ... & Lidow, A. (2023). Projecting GaN HEMTs lifetimes under typical stresses commonly observed in DC-DC converters. Power Electronic Devices and Components, 6, 100051.

[19] Brazzini, T., et al., "Mechanism of hot electron electroluminescence in GaN-based transistors," J. Phys. D: Appl. Phys. 49, 435101, 2016

[20] "Guideline to specify a transient off-state withstand voltage robustness indicator in datasheets for lateral GaN power conversion devices," Dec. 2021. [Online]. Available: https://www.jedec.org/standardsdocuments/docs/jep186

[21] Spirito, P., Breglio, G., d'Alessandro, V., and Rinaldi, N., "Analytical model for thermal instability of low voltage power MOS and S.O.A. in pulse operation," 14th International Symposium on Power Semiconductor Devices & ICS; Santa Fe, NM; 4–7 June 2002; pp. 269–272.

[22] Mishra, S., "Fault current limiting and protection circuit for power electronics used in a modular converter," M.S. thesis, University of Tennessee, Knoxville, TN, 2008. [Online]. Available: https://trace.tennessee.edu/utk_gradthes/468

[23] Glaser, J., "An introduction to Lidar: A look at future developments," IEEE Power Electronics Magazine, March 2017

[24] H. C. Ma et al., "Reliability and failure mechanism of copper pillar joints under current stressing," J. Mater. Sci. Mater. Electron., vol. 26, no. 10, pp. 7690–7697, 2015, doi: 10.1007/s10854-015-3410-8

[25] M. Ding, G. Wang, B. Chao, P. S. Ho, P. Su, T. Uehling, and D. Wontor, "A Study of Electromigration Failure in Pb-Free Solder Joints," Proc 43rd IEEE International Reliability Physics Symposium, San Jose, CA, April. 2005, pp. 518-523

[26] Jae-Woong Nah et al., "Electromigration in Pb-free solder bumps with Cu column as flip chip joints," 56th Electronic Components and Technology Conference 2006, San Diego, CA, USA, 2006, pp. 6, doi: 10.1109/ECTC.2006.1645720

[27] Madanipour, H., Kim, Y., Kim, C., Mishra, D., & Thompson, P. (2021). Study of electromigration in Sn-Ag-Cu micro solder joint with Ni interfacial layer. Journal of Alloys and Compounds, 862, 158043. https://doi.org/10.1016/j.jallcom.2020.158043

[28] N. Islam, G. Kim and K. Kim, "Electromigration for advanced Cu interconnect and the challenges with reduced pitch bumps," 2014 IEEE 64th Electronic Components and Technology Conference (ECTC), Orlando, FL, USA, 2014, pp. 50-55, doi: 10.1109/ECTC.2014.6897266

[29] J. R. Black, "Metallization failures in integrated circuits." doi: 10.1002/nav.3800080206

[30] J. Lienig and M. Thiele, Fundamentals of Electromigration- Aware Integrated Circuit Design. Springer, 2018

[31] Madanipour, H., Kim, Y., Kim, C., Mishra, D., & Thompson, P. (2021). Study of electromigration in Sn-Ag-Cu micro solder joint with Ni interfacial layer. Journal of Alloys and Compounds, 862, 158043. https://doi.org/10.1016/j.jallcom.2020.158043

[32] Guideline for Characterizing Solder Bump Electromigration under Constant Current and Temperature Stress, Version 1.0, JEDEC Standard JEP154,2011

[33] Coonrod, J. (2011). Understanding when to use FR-4 or high frequency laminates. OnBoard Technology, 26-30.

[34] Michaelides, S., & Sitaraman, S. K. (1999). Die cracking and reliable die design for flip-chip assemblies. IEEE Transactions on Advanced Packaging, 22(4), 602-613

[35] Wu, B., Yang, Y. H., Han, B., & Schumacher, J. (2018). Measurement of anisotropic coefficients of thermal expansion of SAC305 solder using surface strains of single grain with arbitrary orientation. Acta Materialia, 156, 196-204

[36] Ghaffarian, R. (2000). Accelerated thermal cycling and failure mechanisms for BGA and CSP assemblies. J. Electron. Packag., 122(4), 335-340.

[37] Han, B., & Guo, Y. (1996). Determination of an effective coefficient of thermal expansion of electronic packaging components: A whole-field approach. IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A, 19(2), 240-247

[38] Barbini, D., & Meilunas, M. (2011). Reliability of lead-free LGAs and BGAs: Effects of solder joint size, cyclic strain and microstructure. SMTA International Proceedings, Fort Worth, Texas, 292

[39] JEDEC Standard, "Temperature Cycling," Test Method JESD22-A104F, November 2020

[40] Cramér, H. (1999). Mathematical methods of statistics (Vol. 43). Princeton University Press

[41] Clech, J. P. (2015, September). Board, package and die thickness effects under thermal cycling conditions. In Proceedings of SMTA International (pp. 40-50)

[42] Tee, T. Y., Ng, H. S., Yap, D., & Zhong, Z. (2003). Comprehensive board-level solder joint reliability modeling and testing of QFN and PowerQFN packages. Microelectronics Reliability, 43(8), 1329-1338

[43] Farooq, M., Goldmann, L., Martin, G., Goldsmith, C., & Bergeron, C. (2003, May). Thermo-mechanical fatigue reliability of Pb-free ceramic ball grid arrays: Experimental data and lifetime prediction modeling. In 53rd Electronic Components and Technology Conference, 2003. (pp. 827-833).

[44] Norris, K. C., & Landzberg, A. H. (1969). Reliability of controlled collapse interconnections. IBM Journal of Research and Development, 13(3), 266-271

[45] Han, B., & Guo, Y. (1996). Determination of an effective coefficient of thermal expansion of electronic packaging components: A whole-field approach. IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A, 19(2), 240-247.

[46] Darveaux, R. (2005). Effect of assembly stiffness and solder properties on thermal cycle acceleration factors. In THERMINIC 2005 (pp. 192-203). TIMA Editions.

[47] Clech, J. (2016). The combined effect of assembly pitch and distance to neutral point on solder joint thermal cycling life. In Proceedings of SMTA International (pp. 25-29).

[48] Motalab, M., Cai, Z., Suhling, J. C., & Lall, P. (2012, May). Determination of Anand constants for SAC solders using stress-strain or creep data. In 13th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (pp. 910-922). IEEE.

[49] Motalab, M., Mustafa, M., Suhling, J. C., Zhang, J., Evans, J., Bozack, M. J., & Lall, P. (2013, May). Correlation of reliability models including aging effects with thermal cycling reliability data. In 2013 IEEE 63rd Electronic Components and Technology Conference (pp. 986-1004). IEEE.

[50] Vasudevan, V., & Fan, X. (2008, May). An acceleration model for lead-free (SAC) solder joint reliability under thermal cycling. In 2008 58th Electronic components and technology conference (pp. 139-145). IEEE.

[51] Sun, F. Q., Liu, J. C., Cao, Z. Q., Li, X. Y., & Jiang, T. M. (2016). Modified Norris–Landzberg model and optimum design of temperature cycling alt. Strength of Materials, 48, 135-145.

[52] Lall, P., Shirgaokar, A., & Arunachalam, D. (2012). Norris–Landzberg acceleration factors and Goldmann constants for SAC305 lead-free electronics. Journal of Electronic Packaging, 134(3), 031008.

[53] Deshpande, A., Jiang, Q., Dasgupta, A., & Becker, U. (2019, May). Fatigue life of joint-scale SAC305 solder specimens in tensile and shear mode. In 2019 18th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm) (pp. 1026-1029). IEEE.

[54] Cui, H. (2005, January). Accelerated temperature cycle test and Coffin-Manson model for electronic packaging. In Annual Reliability and Maintainability Symposium, 2005. Proceedings. (pp. 556-560). IEEE.

[55] JEDEC Standard, "Stress-Test-Driven Qualification of Integrated Circuits," JESD47L, December 2022

[56] Automotive Electronics Council, "Failure Mechanism Based Stress Test Qualification For Discrete Semiconductors In Automotive Applications," AEC-Q101-Rev E, March 2021

[57] Wu, K. C., Lee, C. H., & Chiang, K. N. (2016, May). Characterization of thermal cycling ramp rate and dwell time effects on AF (Acceleration Factor) Estimation. In 2016 IEEE 66th Electronic Components and Technology Conference (ECTC) (pp. 251-256). IEEE.

[58] R. V. Mises, "The mechanics of solids in the plastically-deformable state" (No. NAS 1.15: 88448), 1986.

[59] "EPC2218A – Enhancement-mode power transistor," EPC2218A datasheet, Efficient Power Conversion

[60] "EPC Reliability Report Phase 14." [Online]. Available: https://epc-co.com/epc/design-support/gan-fet-reliability

[61] JEDEC Standard, "Evaluation Procedure For Determining Capability To Bottom Side Board Attach By Full Body Solder Immersion Of Small Surface Mount Solid State Devices," JSED22-A12B, March 2018

[62] Department of Defense Test Method Standard: Mechanical Tests – Die Shear Strength. Mil-Std-883e (Method 2019), May 3, 2018. [Online]. Available: https:// landandmaritimeapps.dla.mil/Downloads/MilSpec/Docs/MIL-STD-883/std883.pdf

[63] AEC-Q200 REV D: Stress Test Qualification for Passive Components (base document), Automotive Electronics Council, June 1, 2010, [Online]. Available: www.aecouncil.com

[64] AEC-Q200-005 Rev A: Board Flex Test, Automotive Electronics Council, June 1, 2010, [Online]. Available: www.aecouncil.com

[65] "$1/W Photovoltaic Systems: white paper to explore a grand challenge for electricity from solar," U.S Department of Energy, August 2010

[66] "SunShot 2030 for Photovoltaics (PV): Envisioning a Low-cost PV Future" by W. Cole, B. Frew, P. Gagnon, J. Richards, Y. Sun, J. Zuboy, M. Woodhouse, and R. Margolis, National Renewable Energy Laboratory, Golden, CO (2017).

[67] "Enphase IQ 7-Based M-Series Microinverters," Enphase Energy data sheet, February 2023

[68] "IQ8M and IQ8A Microinverters," Enphase Energy data sheet, April 2023

[69] "SolarEdge Power Optimizer Module Embedded Solution, OPJ300-LV," SolarEdge data sheet, January 2023

[70] "Power Optimizer Frame-Mounted P370/P401/P404/P500," SolarEdge data sheet, May 2023

[71] "EPC2059 – Enhancement-mode power transistor," EPC2059 datasheet, Efficient Power Conversion

[72] "EPC2218 – Enhancement-mode power transistor," EPC2218 datasheet, Efficient Power Conversion

[73] "EPC2302 – Enhancement-mode power transistor," EPC2302 datasheet, Efficient Power Conversion

[74] MSN weather data for Phoeniz, Arizona

[75] "How to Design Synchronous Buck Converter Using GaN FET Compatible Analog Controllers with Integrated Gate Drivers". [Online]. Available: https://epc-co.com/epc/Portals/0/epc/documents/application-notes/How2AppNote025%20How%20to%20Design%20Synchronous%20Buck%20Converter%20Using%20GaN%20FET.pdf

[76] S. Zhang, S. Gajare, R. Garcia, "Using Test-to-Fail Methodology to Predict How GaN Devices Can Last More than 25 Years in Solar Applications", PCIM Asia 2023 Conference

[77] Efficient Power Conversion Corporation, "EPC9078 – Development Board," [Online]. Available: https://epc-co.com/epc/Products/DemoBoards/EPC9078

[78] Pozo, A., Zhang, S., and Strittmatter, R., "EPC GaN transistor application readiness: phase ten testing," EPC Corp., El Segundo, CA, USA, Reliability Report. Available: https://epc-co.com/epc/design-support/gan-fet-reliability/reliabilityreportphase10

[79] Pozo, A., Zhang, S., and Strittmatter, R., "EPC GaN transistor application readiness: phase twelve testing," EPC Corp., El Segundo, CA, USA, Reliability Report. Available: https://epc-co.com/epc/design-support/gan-fet-reliability/reliabilityreportphase12

[80] Meneghini, M., et al., "GaN-based power devices: Physics, reliability, and perspectives," J. Appl. Phys. 130, 181101, 2021

[81] Efficient Power Conversion Corporation, "EPC23012 – ePower™ Stage IC," EPC23102 datasheet. [Online]. https://epc-co.com/epc/Portals/0/epc/documents/datasheets/EPC23102_datasheet.pdf

[82] De Santi, C. et al, "Review on the degradation of GaN-based lateral power transistors," Advances in Electrical Engineering, Electronics and Energy, Vol. 1, 100018, 2021

[83] Wu, Y. et al., "Activation energy of drain-current degradation in GaN HEMTs under high-power DC stress," Microelectronics Reliability, 54, pp. 2668–2674, 2014

[84] Hu, C. et al., "Investigation of a Simplified Mechanism Model for Prediction of Gallium Nitride Thin Film Growth through Numerical Analysis," Coatings, 7, 43, 2017

[85] Efficient Power Conversion Corporation, "EPC21601 – eToF Laser Driver IC," EPC21601 datasheet. [Online]. Available: https://epc-co.com/epc/Portals/0/epc/documents/datasheets/EPC21601_ datasheet.pdf