Time evolution of off-state degradation of AlGaN/GaN high electron-mobility transistors

M. Montes Bajo1*, H. Sun, M. J. Uren, M. Kuball

Center for Device Thermography and Reliability (CDTR), H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom

The evolution of AlGaN/GaN high electron-mobility transistors under off-state stress conditions is studied by gate leakage current ($I_g$) monitoring, electroluminescence (EL), and atomic force microscope (AFM) imaging at room temperature. It is found that the number of off-state failure sites as identified by EL increases over time during stress until it reaches a saturation value. $I_g$ increases accordingly during stress until this saturation number of failure sites is reached. AFM scanning of the device surface stripped of metal contacts and passivation reveals surface pits corresponding to the location of the EL spots. These pits have an elongated shape oriented towards the drain contact whose length is correlated with the distance to the adjacent pits and with the time since their appearance during the stress experiment. A model for the generation and evolution of the off-state stress-related failure sites is proposed consistent with the experimental results, bringing together surface migration of electrochemical species with trap-based leakage mechanisms and resulting in the formation of an exclusion zone around each failure site.

Keywords: AlGaN/GaN, high electron mobility transistor, electroluminescence, off-state stress, soft breakdown, semiconductor surface, gate leakage current

1 Corresponding authors: Miguel.Montes@icfo.es; Martin.Kuball@bristol.ac.uk

* Present address: ICFO - The Institute of Photonics Sciences, 08860 Castelldefels (Barcelona), Spain.
AlGaN/GaN high electron mobility transistors (HEMTs) show impressive performance in high-frequency and high-power applications. However, the reliability issues of these devices are not yet fully understood and hence attract significant research effort. In particular, the off-state degradation of AlGaN/GaN HEMTs has recently been intensively studied. It is now well known that when a negative bias is applied to the gate electrode of these devices, the gate current, $I_g$, eventually starts to increase progressively. This degradation is accompanied by the appearance of electroluminescence (EL) hot spots next to the drain side edge of the gate metal, and by the physical degradation of the semiconductor in the form of pits and grooves on the device surface. However, it is still not clearly established what the mechanism is that triggers this type of degradation, and also why this degradation takes place in such a spatially localized fashion. Detailed understanding of these mechanisms will be of great benefit to assist in realizing the full potential of AlGaN/GaN HEMTs.

Several mechanisms have been proposed to explain the off-state degradation of AlGaN/GaN HEMTs, such as the generation of trap-assisted percolative leakage paths for electrons in the AlGaN barrier, the structural degradation of the AlGaN barrier due to the inverse piezoelectric effect, the triggering of electric field-driven electrochemical reactions on the semiconductor surface, the diffusion of gate metals into the semiconductor, or the influence of threading dislocations. In this letter, we investigate the off-state degradation of AlGaN/GaN HEMTs focusing on the evolution of the failure sites in terms of their number, spatial distribution, associated EL intensity and gate current, and the morphology of the stress-related semiconductor surface defects. These observations enable the establishment of a model of the generation and evolution of the AlGaN/GaN HEMTs off-state failure sites based on a
combination of percolative electron transport through the barrier and diffusion of electrochemically active species to the leakage site.

The devices studied were AlGaN/GaN HEMTs grown by metal-organic chemical vapor deposition (MOCVD) on SiC substrates, and consisted of a 25 nm-thick AlGaN barrier layer (25% Al) on a GaN buffer layer. Standard NiAu Schottky and TiAlPtAu Ohmic contacts were used for the gate electrode and the source and drain electrodes, respectively. The source-to-gate and source-to-drain distances were 1 µm and 4 µm, respectively, and the gate length was 0.6 µm. The devices were passivated with a multilayer of Si$_3$N$_4$/SiO$_2$/Si$_3$N$_4$ and were isolated by mesa etching. The AlGaN/GaN HEMTs were stressed at room temperature in off-state conditions by keeping a fixed gate bias $V_{gs} = -15$ V (roughly three times the threshold voltage, $V_{th} \approx -5$ V) and a drain-source bias, $V_{ds}$, in the range from 30 to 50 V for periods of time ranging from several hours to several days. Throughout the stress experiment $I_g$ was monitored and EL images were taken using an astronomy-grade charge-coupled device (CCD) camera attached to an optical microscope with a 50× objective with 0.6 numerical aperture.

Figure 1(a) shows results from a 2×100 µm-wide representative device stressed at $V_{gs} = -15$ V and $V_{ds} = 50$ V. At the beginning of the stress experiment $I_g$ decreased as usually observed due to the trapping of electrons underneath the gate contact$^3$ (Figure 1(a) from 1 to 8 s). Soon, $I_g$ started to increase in a stepwise fashion, each step accompanied by the generation of an EL spot along the drain-side edge of the gate (Figure 1(a)). The EL spots, whose appearance did not follow any particular spatial pattern, indicated the device failure sites where gate electrons leak via the AlGaN layer to the 2-dimensional electron gas (2DEG)$^4$.

It might be expected that upon further stress the AlGaN barrier layer will continue degrading, and therefore the generation of EL spots will carry on indefinitely. However,
it is observed (Figure 1(a)) that the number of EL spots increases only until a saturation value is reached after which no further EL spots are generated. Figure 1(b) shows an EL image from one of the fingers of the device after ~4 hours stress, when this saturation had been reached. The stress-related generation of localized failure sites associated with the EL spots is therefore limited, either because the failure sites are linked to material imperfections on the device, and/or because the failure sites have associated an exclusion zone in their immediate surroundings where no additional failure sites can be generated.

The saturation number of failure sites was found not to depend significantly on the bias applied during the stress experiment. After stressing a device at $V_{gs} = -15$ V and $V_{ds} = 40$ V until the number of EL spots was saturated at a value of 234±7, $V_{ds}$ was increased to 60 V and kept at that value for 2 days. Nevertheless, the number of EL spots stayed at 244±8. Therefore, the maximum number of off-state failure sites which emerge in a particular device is determined by the characteristics of the device and the physics of the degradation process themselves, and not by the particular applied electric field.

From Figure 1(a) it is evident that the evolution of $I_g$ essentially mimics that of the number of EL spots. $I_g$ first increases along with the number of EL spots until the saturation value is reached. After this, even though the number of spots remains constant, $I_g$ keeps increasing albeit at a much lower rate, perhaps indicating the presence of an additional degradation mechanism, which will not be discussed further here. From the results in Figure 1(a) it can be estimated that on average each failure site contributes a few µA to the total value of $I_g$. The integrated EL intensity correlates well with $I_g$ and the number of EL spots as it increases up to the saturation point (Figure 1(a)). Interestingly, after saturation of the EL spot generation, the EL intensity decreases
slowly, suggesting that the EL intensity corresponding to each failure site decreases over time, a point which will be addressed in more detail later.

The whole device gate width was mapped with an atomic force microscope (AFM) after removing the passivation layers with an HF:H₂O (1:10) etch, and the metallic contacts with aqua regia (HCl:HNO₃, 3:1) at 80 ºC for 20 min, similar to Ref. 6. The images revealed surface pits adjacent to where the drain-side edge of the gate used to be (Figure 2(a)) which match the location of the EL spots observed during the stress of the devices, as has been reported elsewhere⁴. The pits show some internal structure as if they are composed of smaller units coalesced together, as in Ref. 4. They also appear similar in shape to the pits reported by Gao et al.⁸, identified in their work as the imprint of a stress-generated surface defect which would have been washed out during the chemical removal of the contacts and passivation. These surface defects are in turn identified as the product of an electrochemical oxidation reaction happening on the AlGaN/passivation interface, and driven by the applied off-state stress electric field. The oxygen-related species involved in this surface electrochemical reaction likely have their origin in environmental moisture or the native oxide on the III-N surface. Therefore, it seems likely that the pits from Figure 2(a) could also be the imprint of surface defects produced by a similar electrochemical reaction.

The AFM images show that many of the stress-generated pits have an elongated shape that originates at the gate edge and is oriented towards the drain contact, suggesting that surface defects grow in that direction during stress (Figure 2(a)). Some of the pits, on the contrary, consist of just a small structure next to the gate edge. Typically, pit lengths range from 30 to 400 nm.

The average distance between pairs of adjacent pits is plotted in Figure 2(b) for three different groups of pit pairs: pairs of adjacent pits in which both pits are less than
100 nm in length (short pairs); pairs with both pits longer than 150 nm (long pairs), and a third group including the rest of the pairs (mixed pairs). It is found that the distance between adjacent pits is correlated to their length, indicating that the growth of a pit is affected by the presence of neighboring pits. This would be consistent with a scenario in which each failure site apparent in EL, i.e. each pit, is surrounded by an exclusion zone in which no more pits can be generated, rather than the alternative scenario of pits being associated with pre-existing defects. The size of such an exclusion zone correlates with the size of the pit, either because pits with a larger initial exclusion zone can grow longer, or because the size of the exclusion zone grows with the size of the pit.

To help to build a model of the evolution of the failure sites, it is useful to correlate each pit’s length towards the drain with the time elapsed between the generation of the corresponding EL spot and the end of the stress experiment (i.e. the pit’s age, Figure 3(a)). During this time the device was kept under stress and a leakage current was flowing through the failure sites, so the stress-related surface defects would have had the opportunity to grow. It is found that pits of any particular age can reach any length up to a maximum. The later the pit forms during stress, the shorter the maximum attainable pit length is, i.e. ‘old’ pits can be of any length whereas ‘young’ pits can only be short. Therefore the pits can grow with time, but finding ‘old’ pits with a wide range of lengths also indicates their growth can be stopped for some reason. One possibility is that the growth of one particular pit is halted because during device stress it may end up in the exclusion zone of a neighboring pit, thus preventing its further evolution.

It could be argued that younger pits are shorter because they have not had enough time to grow, i.e. if the stress experiment had progressed longer, the corresponding surface defect would have grown longer. However, Figure 3(a) shows the difference in age between the youngest and the oldest pits is less than $20 \times 10^3$ seconds, whereas the
total stress experiment lasted for over $600 \times 10^3$ seconds, i.e. young and old pits have comparable times of total stress. Note that the saturation of the EL spot generation occurred at around $20 \times 10^3$ seconds into the stress experiment on this device. Therefore, it appears that the growth of the failure sites progresses only until the saturation point is reached.

Figure 3(b) shows the integrated EL intensity as a function of stress time during the whole stress experiment for three cases of the failure sites represented in Figure 3(a). The fluctuations in EL intensity indicate corresponding fluctuations in the leakage current flowing through the failure site. It can be seen that following saturation the fluctuations become more stable and the EL intensity undergoes a gradual reduction consistent with the decrease of total EL intensity after saturation shown in Figure 1(a).

Using the results shown, a model for the generation and evolution of the device failure sites during off-state stress can now be built. When the stress experiment begins there are no clearly visible EL spots or localized leakage paths (Figure 4(a)), although there is a finite initial $I_g$ as seen in Figure 1(a). This current is probably due to tunneling of electrons through the AlGaN barrier aided by the high electric field at the edge of the gate and the possible presence of traps in the barrier\textsuperscript{15}. The initial $I_g$ is most probably not homogeneously distributed along the gate contact edge, and shows local maxima due to inhomogeneities in the electric field, related to gate-edge roughness, or crystallographic features of the AlGaN barrier, such as clusters of defects, dislocations\textsuperscript{16}, or step edges\textsuperscript{17}. Degradation would likely start at those sites, assuming that the surface pits form due to an electrochemical reaction driven by the local electric field\textsuperscript{8}. A conductive defect would be produced on the semiconductor surface (the imprint of which is observed by AFM, Figure 2(a)), followed by localized breakdown
via a percolative path through the AlGaN barrier with a current limit of a few µA (Figure 4(b)).

The reduced resistance between device top surface and 2DEG would lower the electric field at the surface in the area surrounding each failure site, forming a local exclusion zone for the electrochemical reaction. However this would not be sufficient to explain the large exclusion region extending up to a micron surrounding each defect site seen in Figure 2, nor the fact that its size is not affected significantly by applied bias. A plausible explanation is that the formation of the surface defects requires the field-driven migration of electrochemically active positively charged species across the surface of the AlGaN to the failure site where they react with interfacial species, possibly the oxygen-related species of Ref. 8. The mobile species themselves are unlikely to be oxygen containing given its high electronegativity. The supply of these positively charged species would limit the electrochemical reaction, with a denuded zone forming around each high field region at surface defect tips (Figure 4(b)). The volume of material involved in the surface defects, which are only a few nm deep at most, is dramatically smaller than would be expected if the leakage current through the failure site were carried by an electrochemical process thus consuming material. Hence there are two parallel processes in operation, an electronic transport mechanism, responsible for the vast majority of the current shown in Figure 1(a), and an electrochemical process responsible for the consumption of material and formation of the surface defects. Due to the small radius of curvature of each electrochemically originated surface defect, the electric field at its tip would be enhanced and would attract the migration of charged chemical species allowing the surface defect to grow, thus lowering its curvature and slowing its growth. The formation of further high curvature locations allowing the surface defect growth to restart (as seen in Figure 2(a))
would require fluctuations in the tip field perhaps driven simply by stochastic processes (Figure 4(c)). Additional contributions from non-stress related, pre-existing defects may also be possible.

The EL plot in Figure 3(b) demonstrates that the light emission fluctuates but is not proportional to the pit length, consistent with the leakage path between the surface defect and the 2DEG through the barrier being highly localized, and without multiple local leakage paths except possibly transitorily. However, does this local leakage path migrate along the surface defect as it grows? There is no direct evidence for this process, but the reduction in noise in Figure 2(b) after saturation might suggest that the vertical leakage path can change location before saturation occurs.

When the density of failure sites increases, the exclusion zones of denuded chemical species surrounding the surface defects overlap (Figure 4(d)) and eventually all the species are consumed and the generation of further failure sites or the growth of existing ones stops, as demonstrated in Figure 1(a). This also provides a natural explanation for the correlation between pit spacing and length seen in Figure 2.

In conclusion, off-state degradation of AlGaN/GaN HEMTs has been studied with a focus on the evolution of device failure sites. The stressed devices fail on a finite maximum number of sites intrinsic to the device and not affected by the applied electric field. AFM imaging of the devices stripped of contacts and passivation revealed the failure sites have associated pits on the semiconductor surface featuring an elongated shape pointing towards the drain contact. Pit length is correlated with the distance to their neighbors. Also, pits formed early in the stress experiment are found of any length up to a maximum, whereas pits formed at later stages are always short. With these results, a model of the generation and evolution of the device failure sites has been developed. Failure sites would be generated earlier at sites where the local electric field,
and therefore $I_p$, is higher, maybe linked to defects on the semiconductor or fabrication imperfections. An exclusion zone appears around each failure site, where the generation of additional failure sites is precluded. It is proposed that this exclusion zone is the result of the consumption of mobile, positively charged chemical species in an electrochemical reaction\(^8\) resulting in surface defects. As the stress experiment progresses, the surface defect and its corresponding exclusion zone grow. The process of generation of EL spots during off-state stress saturates when the exclusion zones from all the failure sites overlap along the whole width of the device and the chemical species are all consumed.

The authors would like to thank A. Murray (University of Bristol) for his help on the removal of the contacts and passivation from the devices. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under Grants EP/H011366/1, EP/K014471/1, EP/K024345/1 and EP/L007010/1. The HEMTs were fabricated at QinetiQ Ltd.

---

Figure 1. (a) Gate leakage current, number of EL spots, and integrated EL intensity as a function of stress time for a 2×100 µm-wide AlGaN/GaN HEMT device stressed at $V_{gs} = -15$ V and $V_{ds} = 50$ V. (b) False color EL image overlaid on a white-light image from one of the fingers of the device in (a) after $\sim10^4$ seconds stress, i.e. after the saturation of the EL spot generation.
Figure 2. (a) Atomic force microscopy (AFM) image of the region near the gate of an AlGaN HEMT device after removal of the contacts and passivation. The white dotted line indicates the original location of the drain-side edge of the gate. For this device, the saturation value of off-state stress failure sites had been reached. (b) Average distance between pairs of adjacent pits calculated for three groups: Short pairs, in which both adjacent pits are less than 100 nm long; long pairs, in which both pits are more than 150 nm long; and mixed pairs, for any other case. Error bars indicate the standard deviation from the average.
Figure 3. (a) Length of the surface pits towards the drain as a function of the corresponding EL spot onset time and the pit age (i.e. the time between the appearance of the EL spot and the end of the stress experiment). (b) EL integrated intensity as a function of stress time for three representative EL spots from (a): A ‘young’ EL spot associated with a short surface pit, an ‘old’ EL spot associated with a long pit, and an ‘old’ EL spot associated with a short pit. The dotted line indicates the approximate time of saturation of EL spot generation.
Figure 4. Schematic of the proposed model for the evolution of the EL spots. (a) Fresh device. Initial $I_g$ is most likely not homogeneously distributed along the width of the gate. (b) A failure site with its associated surface defect has been generated, most likely at a location with higher initial $I_g$. The surface electrochemical reaction generates an exclusion zone (dotted line) denuded of mobile, positively charged species that have migrated to the surface defect. (c) As the stress progresses further the surface defect and its exclusion zone grow. Other surface defects with their corresponding exclusion zones are also generated. (d) The EL spot saturation number is reached when the exclusion zones from all the failure sites overlap along the whole gate. The pits cease growing when all the mobile species are consumed.