Low-cycle fatigue of multilayer metal stack employed as fast wafer level monitor for backend integrity in smart power technologies

Alexander Mann[a,b,*], Henning Lohmeyer[a], Yvonne Joseph[b]
[a]Robert Bosch GmbH, Automotive Electronics, Tübinger Straße 123, 72762 Reutlingen, Germany
[b]Technische Universität Bergakademie Freiberg, Institute of Electronic and Sensor Materials, Gustav-Zeuner-Straße 3, 09599 Freiberg, Germany
*email: alexander.mann@de.bosch.com

Abstract

A novel approach for wafer-level test and monitoring of multilayer metal-stack integrity in integrated circuit process technology based on the low-cycle fatigue of power device metallization structure is described. Repetitive power pulsing at the limit of the electro-thermal safe-operating area of the devices reveals systematic changes in level and homogeneity of intrinsic thermomechanical robustness and is able to activate latent defects. Exemplarily for two smart-power process technologies the intrinsic low-cycle lifetime limit is explored as reference basis and transfer to test vehicles on product or process control module is validated in experimental case-study and supported by detailed electro-thermal simulation of stress pulse events.

1. Introduction

Reliability of multilayer metallization stack is crucial for highly reliable integrated circuits (IC) used for automotive and industrial applications. Common methods like high temperature bake on wafer or temperature cycling of packaged devices allow for qualification of thermomechanical stability (stress induced voiding) and robustness (chip-package interaction) but are not suitable as fast tests to monitor mechanical integrity of on-chip backend stack. We propose to use a highly accelerated repetitive power pulsing test (RPP) on integrated power stages on wafer level to monitor mechanical integrity of smart power ICs.

The procedure is derived from applications using power stages as low side switches for control of inductive loads, e.g. magnetic valves. Each switching cycle can involve a millisecond power pulse leading to a pronounced local temperature excursion in the driver stage. A strong lateral temperature gradient imposes severe stress on the IC backend structures due to the mismatch of the coefficient of thermal expansion of aluminum or copper metals and dielectrics which may cause degradation and IC failure in respective applications [1,2,3,4,5]. Compared to vertical discrete power devices, where cyclic thermomechanical stress leads to pronounced aging of the usually thicker top aluminum metal layer itself [6], for smart-power IC products additional failure channels are present related to dielectrics and various interfaces in the fine-combed multi-finger metal stack. Typically for product application the challenge is to guarantee safe operating life for a given mission profile which may involve billions of cycles [7]. In contrast for our purpose we explore the low-cycle lifetime limit for devices under RPP stress at the limit of the single-pulse safe-operating area (SOA) of the drivers given by electro-thermal runaway [8].

In the first section of the paper we describe the experimental setup used together with an approach for electro-thermal simulation which gives us detailed insight on temperature and current density distribution in the devices under test (DUT). By executing systematic end-of-life trials afterwards we explore the intrinsic low-cycle fatigue regime for two different smart-power technologies. Based on these results the feasibility of using integrated power stages as wafer level reliability (WLR) monitor for backend integrity is discussed and a case study is presented. In the last section area scaling of RPP stress is investigated to validate the possibility of embedding dedicated test vehicle in the process control module (PCM) for monitoring purposes.

2. Experimental procedure and simulation approach

As a test vehicle in principle any integrated high-voltage (HV) device is usable, given it can be scaled in area and can operate at points of high power dissipation, i.e. with elevated voltage drop over the device. We employ lateral n-channel metal-oxide semiconductor field effect transistors of voltage class 40V or 60V maximum drain-source voltage typically available in current smart-power technologies to be used for low-ohmic low-side switches. The devices are layouted with multi-finger geometry and an active area of typically 0.3mm² unless stated otherwise. We investigate two different smart-power technologies both having a four level metallization stack consisting of three aluminum levels and one top copper level. Conveniently the DUT is available with open drain and rectangular current pulses are applied via a drain-gate clamp circuit, see e.g. setup described by Kendrick et al. [9]. In this case the clamp to control the MOS gate consist of a chain of Zener diodes between drain and gate and a gate pull-down resistor between gate and source and can easily be integrated together with the DUT on wafer. Alternatively often suitable low-side channels are integrated in respective IC products such that our test method can be implemented directly on production wafers. A Keithley 2430 1kW Pulsed Mode SourceMeter is used to apply the rectangular power pulses in an automated setup for WLR tests on a semiautomatic probe station. An end-of-life trial for a DUT involves bursts of millisecond pulses with repetition rate of typically 100Hz of increasing length. Periodically intermediate leakage and $R_{DS(on)}$ measurements on the DUT are executed. DUT failure is detected if leakage current or $R_{DS(on)}$ increases above a

978-1-5090-2106-2/16/$31.00 ©2016 IEEE
DOI: 10.1109/EuroSimE.2016.7463315

certain threshold value or even clamp voltage during pulsing collapses due to short of DUT.

To obtain fast feedback in a WLR monitor test and large failure statistics short test times for a single DUT are desirable. In our context this means a high maximum temperature at end of pulse together with a high temperature swing per pulse should be reached, respectively. At the same time operation within the electro-thermal SOA of the MOS device has to be guaranteed to avoid failures which are not related to the mechanical stress. For characterization purposes we use measurements on dedicated test structures with multiple tightly embedded NPN-type temperature sensors to determine stress conditions precisely [10]. For the calibration of these NPN-sensors a thermochuck from Temptronic is used to perform a DC characterization up to 300°C. Based on this data the temperature T can be modeled regarding the thermovoltage V_{BE} of the NPN-Structure using an Ebers-Moll-approach

$$V_{BE} = \frac{n_F kT}{q} ln\left(\frac{I_E}{I_0} + 1\right) \text{ with } I_0 = cT^3 exp\left(\frac{-E_g}{kT}\right).$$

In this equation, V_{BE} represents the base-emitter voltage and I_E the emitter current of the NPN-structure. The ideality factor n_F, the silicon bandgap E_g of the highly doped emitter region, and c are used as fitting parameters. Validity of the extrapolation of this model approach is determined in the literature up to 600°C [10].

In addition to directly measure the transient hot-spot temperature, pulse events have been analyzed with electro-thermal simulation. This method shows local current density, power density and temperature distribution in the devices including the multi-finger metal stack in full detail. By this approach transfer of test conditions to test structures without sensor, e.g. output channels of given products, is possible. For simulation the electro-thermal analyzer ETHAN from Silicon Frontline Technology Inc. (SFT) has been used. ETHAN is based on electrical modeler R3D from SFT [11], and allows self-consistent transient solution of current flow in the metals and power dissipation in the silicon active part of the HV-MOS device based on local operation point and temperature. Thus electro-thermal feedback within the device is included in the transient simulation. Local device behavior is simulated based on a Spice-like electrical device model of the HV-MOS, thus high accuracy of the latter at elevated temperature is a prerequisite.

Figure 1 shows the measured transfer characteristics of a small reference MOS-device of type which will be used in RPP tests below and results from evaluation of its Spice-model for ambient temperature of 25°C, 150°C and 300°C. Model and measurement are in very good agreement, in particular for the relevant region of small gate overdrive (V_G < 2V) and deviation at high temperature is below 5%. Thus the precision of the electrical model at high temperatures is sufficient and a physically plausible extrapolation up to 300°C has been checked. Since ETHAN determines the local power density in a distributed manner based on the Spice model and the electrical stimuli, device internal electro-thermal feedback

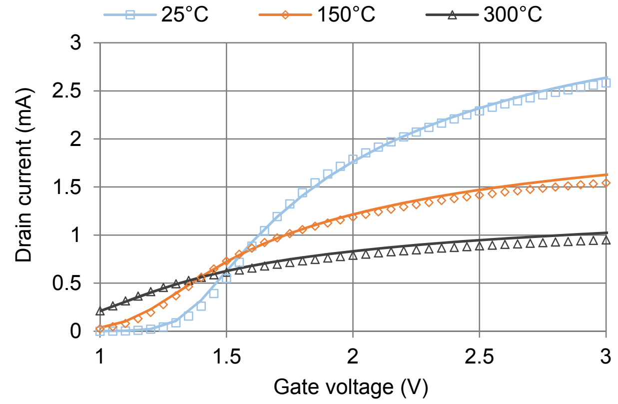

Figure 1: Transfer characteristic of a small reference MOS-structure for the chuck temperatures 25°C, 150°C and 300°C. The Dots represents the measured data points and the lines represents the simulated curves based on circuit simulation.

will be reproduced, which will be positive when working at small gate voltage in saturation (for the present device for V_G below 1.5V). For very high simulated temperature device model values are clipped by the simulator to 400°C. So we do not expect to directly reproduce the ultimate electro-thermal runaway of the device with the current simulation setup [12].

For a single pulse representative for the RPP trials Figure 2 shows the transient temperature in the central hottest region of device both measured and simulated. A very good agreement between simulated and measured temperature is obtained. The inset shows the lateral temperature distribution within the device at the silicon surface which is introducing the severe lateral mechanical stress gradients in the multilayer metal stack responsible for degradation under RPP stress. The simulation setup is

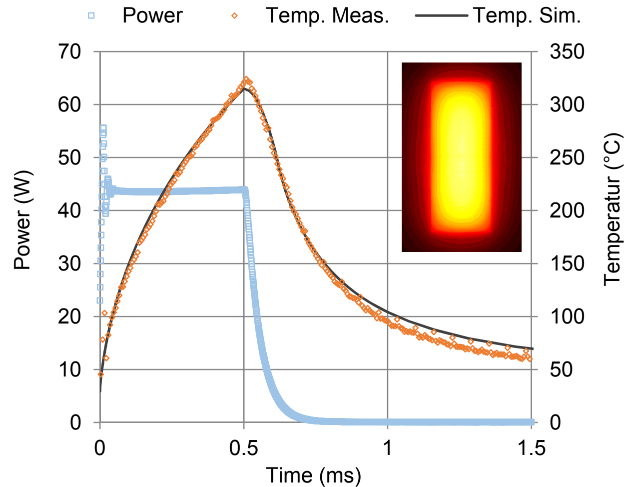

Figure 2: Rectangular power pulse and measured hot spot temperature for MOS test structure with embedded temperature sensor in comparison with electro-thermal simulation result. The inset shows the temperature map of the silicon active area at the end of the power pulse (t=0.5ms; scale: 25°C (dark) – 325°C (light)).

directly based on the original layout data of the integrated power MOS together with the technology layer stack description and the device model from the process development kit. Since experiments under consideration are done on wafer level modeling of heat conduction path from device active area through silicon substrate to wafer backside is sufficient to obtain accurate results. Here thermal conductivity of doped substrate has been adjusted and lateral simulation domain is large enough to allow heat spreading in the silicon substrate. The backside of the wafer is thermally fastened to the 7.0mm thick thermochuck top-plate. Assuming the material of the chuck is pure copper, the area-specific thermal conductivity and the heat capacity accounts to 8.9E-11 W/µm²/K, and 2.4E-8 J/µm²/K, respectively, which are used as wafer backside boundary condition in the simulation.

Validated through the transient temperature measurements electro-thermal simulation thus allows aggressively pushing repetitive pulsing tests to the limits of electro-thermal SOA, i.e. to values of temperature swing well above 350K. This allows investigation of a novel fast wafer level monitor method.

3. Intrinsic failure regime of low-cycle RPP-stress

To show the possibilities of using a MOS-structure of the type described as backend integrity monitor in WLR, we discuss the method by means of a technology comparison. Both processes are smart power technologies with a backend metallization stack consisting of three metal layers aluminum and one top metal-layer of thick copper. As Table 1 indicates, the important difference of these two technologies -referred here by A and B- in this context is a variation of the backend metal layer thickness.

Table 1: Normalized thickness variation of the backend metal layers of two different smart power technologies. The first three metals and the top metal consists of aluminum respectively copper.

Thickness (nom)	Technology A	Technology B
Metal 1	1	1.08
Metal 2	1	1.22
Metal 3	1	1.95
Metal Top	1	1.26

For both technologies RPP test series were executed under accelerated conditions. Length of rectangular power pulses has been kept constant to 0.5ms while power (i.e. drain current at constant drain-source voltage) were varied. For technology B additionally trials at different ambient chip temperatures T_{chip} were collected.

Figure 3 summarizes the lifetime results with number of pulses to failure being the 63% percentile of the Weibull distribution fit of the end-of-life data obtained for both technologies. As expected with increasing thermal swing of the pulse event, the mean lifetime, i.e. the t63 value decreases strongly, in particular for technology B. Here the RPP failure mechanism active in a low-cycle fatigue regime may be assumed as follows: The large thermal swing leads to high lateral mechanical stress states beyond the yield point of the metal thin films such that each cycle may accumulate a viscoplastic deformation of the metal. Ultimately stress on dielectrics may exceed critical limits and an electrical detectable leakage path or even short in the drain/source metal path of the driver results [3,4,5]. We checked by failure analysis, that DUT failure indeed is caused by shorts in metal system and is unrelated to Si active device.

Typically this regime may well be modelled through Coffin-Manson like approach [2,9] or extended models like Norris-Landzberg ansatz [13]. We find that a modified Norris-Landzberg based ansatz, typically used for a low-cycle fatigue based failure mechanism, fits mean number of pulses to failure N_f well for both technologies:

$$N_f = A \cdot (\Delta T)^{-n} \cdot exp\left(\frac{E_a}{k_B T_m}\right) \text{ with } T_m = T_{chip} + \frac{\Delta T}{2} ,$$

where ΔT and T_{chip} represent the thermal swing and the ambient temperature, respectively. The technology-dependent parameter A is used as a fit variable as well as the Coffin-Manson exponent n, which we determine in the range of 8 to 11 indicating a mechanism related to a brittle fracture. For technology A the Arrhenius term is set to 1 since no data for varied T_{chip} is available.

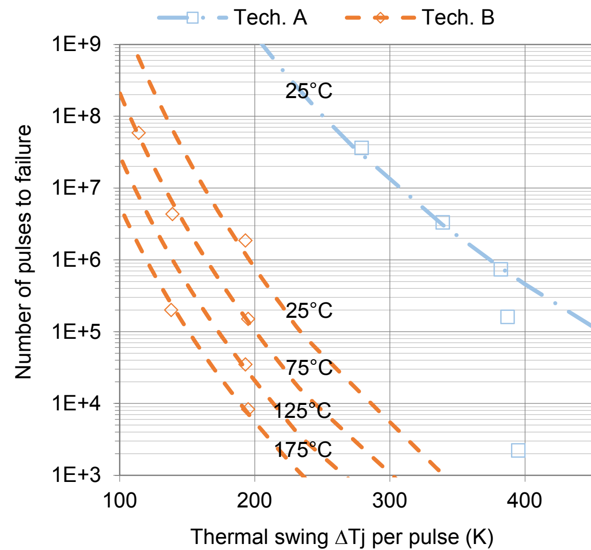

Figure 3: Coffin-Manson chart of mean lifetime (t63 value) based on measured test series for technology A (blue dots) and B (orange dots) together with appropriate model (dashed blue and orange lines). Numbers indicate here T_{chip} of the respective trails.

The dashed lines in Figure 3 represent the respective model fit for both technologies. In case of technology B data from package level trials with longer lifetime has been included to demonstrate the good extrapolation capability of the simple approach. The data indicates a strong difference of the number of pulses to failure for both technologies. At ambient temperature of 25°C and comparable stress conditions related to the same thermal swing of about 200K, the technology A can withstand about three orders of magnitudes more pulses until the

failure occurs. We attribute this difference to the influence of the metal thickness in particular of metal 3 of both technologies. As has been discussed by Smorodin et al. metal thickness has a pronounced influence on the yield stress value of the thin Aluminum films [14]. So generally a smaller layer thickness of the aluminum layers is desirable if the IC has to withstand repetitive thermal cycling. Here in addition changes from B to A in the layout topology of the uppermost thin metal layer are found to have a beneficial effect on RPP robustness of the final MOS device.

In case of technology A the Coffin-Manson model fits well up to a thermal swing of about 380K. Further increasing pulse energy obviously results in a change of main failure mechanism for these devices of typical area. The mean number of pulses to failure collapses and at the same time the failure distribution strongly widens. Typically the trials exhibit quite tight failure distributions with Weibull slope above 2 (compare case study for B below). But for the trials with ΔT above 380K Weibull slope decreases to values below 1. We discuss electro-thermal instability and electromigration as possible additional failure mechanism in this regime.

Since we are working at a highly accelerated stress condition the deviation from Coffin-Manson trend at about ΔT=382K may be due to electro-thermal runaway if we exceed limits of SOA. To check this possibility we measured the thermal runaway of the transistor during a single pulse event using a reference structure with build-in temperature sensors. As shown in Figure 4 we use the stress condition of 58.8W which generates a thermal swing of 339K for 0.5ms and extend the pulse length up to 1.0ms. After a pulse duration of 0.81ms the power curve collapses and the measured temperature shows a steep increase up to 650°C indicating the local thermal runaway at hot-spot of the device. A permanent damage of the device and backend metallization stack occurs. As has been pointed out in the previous section present ETHAN simulation model does not include this runaway indicated by an inflection point in the thermal transient. Below this point, however, up to about 480°C central device temperature, we can both simulate T_{max} reasonably well with current ETHAN setup and in experiment we can switch off the device safely again without destruction.

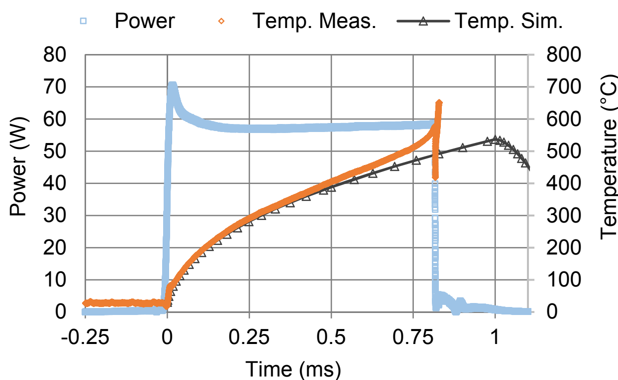

Figure 4: Power plotted together with measured and simulated temperature for a typical MOS-structure based on a stress pulse of 58W for 1.0ms. At t=0.81ms device is destroyed by local thermal runaway.

This result indicates a very good electro-thermal stability of the device and we thus conclude that the observed collapse of the pulses to failure for technology A under high accelerated stress conditions isn't a result of the violation of thermal SOA of the device itself. This is consistent with the observation that the respective trials do not contain events of failure at the first (few) pulses.

In principle electromigration phenomena on the metal lines and VIA structures have to be excluded from influencing the lifetime under these RPP trials considering the high peak temperature involved [2]. To check, we evaluate the peak current density and therewith the weakest spot of the structure. The ETHAN simulation result indicates a small area of the third metal level with a maximum current density of about 3mA/µm² for the given stress condition, see Figure 5. This region is responsible for distribution of the current to areas of the MOS which are not directly connected to an upper copper plate. At the maximum temperature this current density corresponds to a maximum power density of Joule heating in the third metal level of 3.8E-07W/µm³. This is small compared to the power density of about 8.4E-04W/µm³ present in the active device at Si surface. Thus there is no significant

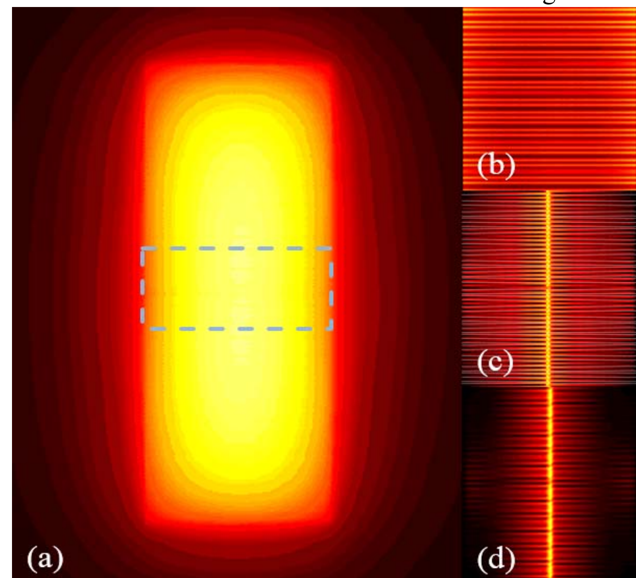

Figure 5: Maps of ETHAN field results simulated for a 70.4W power pulse plotted at t=0.5ms: (a) Lateral temperature distribution of the third metal level of the metallization stack (scale: 50°C (dark) – 500°C (light)). (b) Local power density in the active area of the MOS-device. A small inhomogeneity in the multi-finger structure can be determined due operation conditions (scale: 0.6mW/µm³ (dark) – 0.9mW/µm³ (light)). (c) Local current density (scale: 0mA/µm² (dark) – 4mA/µm² (light)) and (d) local power density (scale: 0µW/µm³ (dark) – 0.4mW/µm³ (light)) of metal fingers in the third metallization level. In comparison to (b) the power density (d) of the metal is scaled with the factor 1000.

additional temperature spread between Si surface and the metal lines, and otherwise such weak spots would show up in ETHAN simulation.

Considering the peak current density in the hot spot region we estimate the expected lifetime based on standard electromigration model following Black´s equation:

$$MTTF = A \cdot J^{-n} exp\left(\frac{E_A}{k_B T}\right),$$

where the mean time to failure $MTTF$ is based on a technology-dependent parameter A, the current density J, a scaling factor n, and an Arrhenius term which consist of the activation Energy E_A and the absolute temperature T. For the known parameters n and E_A of the respective used technology A, the equation results in a sub 1% failure rate after about 4000s cumulated operation time for a maximum current density of 3mA/µm² at an absolute maximum temperature of 690K. This time corresponds to a number of pulses to failure of about 8.2E+06 considering the duty cycle. Thus even for these harsh conditions straightforward conservative assessment of quasi-dc electromigration effects for the aluminum layers at the worst case position and temperature still leaves much headroom and cannot explain the drop in RPP lifetime seen for the typical structure at thermal swing of about 398K. Further considering Black´s equation as a general trend we would expect a steady progression of the lifetimes for temperatures above 400K, too. Possibly failure behavior may be explained by accelerated aging effects of the AlCu metals, i.e. large changes in local resistance of AlCu layer related to the thermal cycling at very high temperatures as discussed by Ferrara et al. [15], which in turn may yield device failure due to local current density hot-spots.

4. RPP WLR monitor test – case study

Practical application of this RPP test methodology on wafer based on intrinsic low-cycle fatigue as monitor for mechanical backend integrity is limited by test time for a single DUT since options for parallel test on wafer are limited and fast feedback is desirable. Considering that repetition rate will not be much higher than 200Hz in practice for constraints of thermal dissipation RPP test till end-of-life will not be an option for technology A, which shows a number of pulses to failure above one million also for a thermal swing of 400K. For technology A -however- intrinsic wear-out point according to the Coffin-Manson data goes down to several thousand pulses for elevated thermal swing.

As a case study to demonstrate the usefulness of the new test approach, the results of trials for a process variation with an influence on the backend robustness of technology B is shown in Figure 6. The tests were executed on wafer level on product output drivers under highly accelerated stress conditions which results in a thermal swing of 270K at an ambient temperature of 25°C. For these conditions RPP stress degradation leads to a number of pulses to failure below 5E+04. In comparison to material from a reference lot B1, a wafer from lot B2 is tested, whose processing included changes in chemical-mechanical planarization (CMP) procedure.

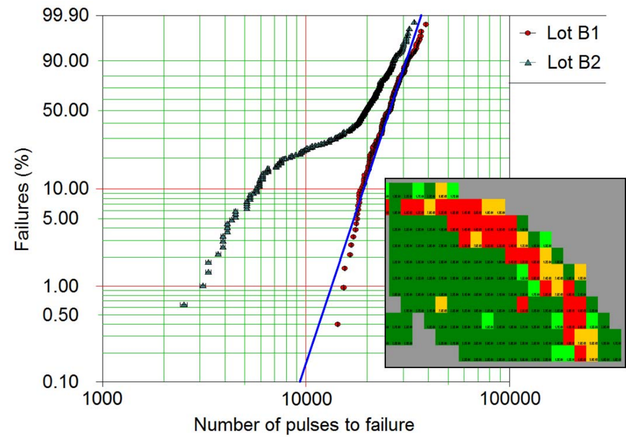

Figure 6: Lifetime distribution as result of wafer level RPP test for two lots B1 (reference) and B2 (subject to process variation) of same technology B, quarter of wafer each. The blue line gives a Weibull fit to lot B1 result with mean lifetime of 27000 pulses and steep slope=6.5. Lot B2 exhibits a bimodal distribution. The inset shows failure data of lot B2 as color-coded wafer map: dark green – DUT lifetime > 17000 pulses, red – DUT lifetime < 7000 pulses.

The RPP test on the reference lot B1 yields a steep Weibull distribution (Weibull slope β=6.5) with a mean lifetime of 27000 pulses which meets expectation from Coffin-Manson modelling. In contrast the distribution of the measured data of lot B2 exhibits a bimodal behavior with higher lifetime branch meeting the reference distribution of lot B1. The inset of Figure 6 shows the failure data of lot B2 as a color-coded wafer map. The map reveals a middle to edge inhomogeneity of the robustness since all devices from the lower lifetime branch are located on a shell near wafer edge.

Obviously in this case changes in the CMP process module did influence the intrinsic RPP robustness of the material which shows up as a systematic trend in the RPP WLR test: For the dies located near wafer edge number of pulses to failure decreases about a factor of two. So usage of the RPP WLR methodology allows fast feedback to technology development and fabrication on robustness of backend stack and on possible systematic influence of changes in process flow.

Additionally defect events or latently weak interfaces may show up in the trials as an extrinsic early-failure branch in the Weibull distribution of the RPP WLR trial (not shown). To detect these effects the test may be implemented as a pass/fail test with suspending all devices not yet failed at a certain (low) pulse number. In this configuration it is applicable for a case of technology A, too, where end-of-life testing is not possible due to high intrinsic robustness.

5. Area scaling of RPP stress degradation and test in process control module

In the previous section we demonstrated usefulness of an RPP end-of-life test on wafer-level allowing test with high statistics and high area coverage with fast turn-around

time. For monitoring purposes however RPP test in PCM would be highly useful as well since test on regular basis during manufacturing would be possible. For such an approach, in addition to harsh constraints on test time, area of DUT has to fit into typical PCM modules, e.g. size in one dimension should be (much) below 100µm. So in the following we investigate area scaling of the RPP failure mechanism for our regime of low-cycle fatigue to assess portability of wear-out mechanism. For these tests integrated power stages of technology A and B with much smaller active area of 0,009mm² and 0,045mm², respectively, have been designed and investigated.

Regarding the main aging mechanism it is established that the degradation of Al metal system under RPP stress is driven by thermal gradients [3] which means that different temperature profiles in a structure may result in strongly different lifetime even for the same maximum temperature [16]. In our case, because of the smaller lateral width of the structures intended for PCM, the lateral temperature profile is flatter internally, i.e. over the lateral extend of the metallization stack, compared to a larger typical MOS power stage as investigated previously. This is illustrated in Figure 7: Here for equal maximum temperature at end of power pulse internal lateral temperature difference measured from the middle to the edge of a large and a small DUT accounts to about 150K and 50K, respectively. Additionally, as shown in Figure 8, for the small DUT the hot spot temperature rises much faster in the first 100µs of the pulse event and runs in a quasi-plateau state, whereas for the larger DUT the hot-spot temperature increases slower. Both differences are caused by lateral heat conduction effects which causes larger internal temperature difference in the case of the larger DUT and a relatively larger external Si volume being heated up by the smaller DUT.

Based on these general differences in transient behavior and lateral distribution of temperature significant differences in RPP degradation behavior may be expected comparing larger structures and the small ones intended for

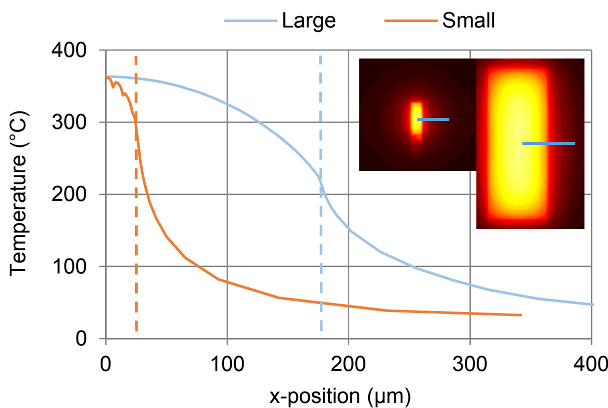

Figure 7: Lateral temperature profile for a large and a small MOS-structure at the end of a pulse event with ΔT=370K. Cutlines are indicated in the insets, x=0 corresponds to device center. The dashed lines indicate edge of active device.

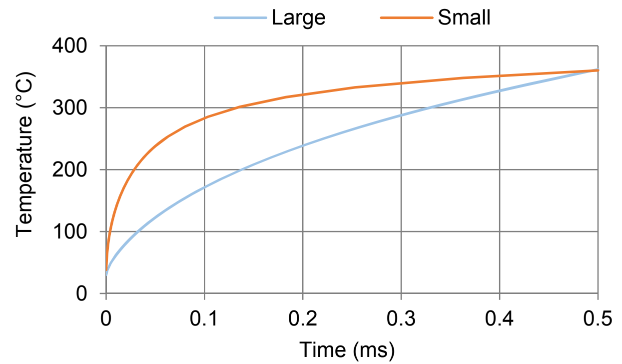

Figure 8: Transient temperature curve for a large and a small MOS-structure simulated for the hot spot of the device under accelerated test condition.

PCM, even if geometrically the layout topology and typical dimension e.g. of metals spacing for the small DUTs is kept as similar as possible to the larger ones.

However in end-of-life trials again focusing on the low-cycle fatigue limit generally we find a very good fit of lifetime of the smaller DUTs to the Coffin-Manson trends as discussed in section 3. Figure 9 summarizes these investigations showing a Coffin-Manson plot with focus on the more robust technology A. For technology B here we report only a single trial on the small DUT which fits very well to the general RPP lifetime model based on Norris-Landzberg approach as discussed before.

For technology A the RPP lifetime for the trials on the small DUT is in agreement with the general Coffin-Manson trend obtained for the larger DUTs for the moderate stress conditions up to ΔT=400K. For even higher stress condition interestingly the RPP lifetime still follows the Coffin-Manson trend up to ΔT=470K. Afterwards for nominal stress condition of ΔT=500K, as simulated with ETHAN, failure distribution broadens and individual DUTs fail at single pulse, indicating limit of device electro-thermal SOA as discussed before.

This means for the small DUTs of technology A we do not observe the abrupt change in RPP failure behavior with sporadic occurrence of much earlier failure which has been observed in the trials on the larger power stage structures. This might be attributed to slightly lower values of maximum current density present in the lower metals of the small DUTs if current flow is assumed to be relevant for the additional failure mechanism active in the larger structures under the present harsh conditions (see discussion under section 3).

Here we conclude with the observation that for both technologies RPP WLR tests on small enough DUTs are feasible which may be integrated in PCM module (e.g. scribe line) and which nevertheless reproduce the intrinsic low-cycle wear-out trend under RPP stress. For technology A this holds true up to the electro-thermal SOA of the device while for technology B intrinsic lifetime is already very low when going to very high maximum temperature. Generally regarding design of dedicated test structures for thermomechanical robustness in PCM alternatively usage

of small backend-stack structures passively heated by on-chip poly-resistors may be a viable approach for IC technologies, too. By using these methods typically employed in fast-WLR procedure [17,18] transfer of our approach to IC process technologies not offering high-voltage power devices should be possible without the need for excessively high current levels in wafer-level setup.

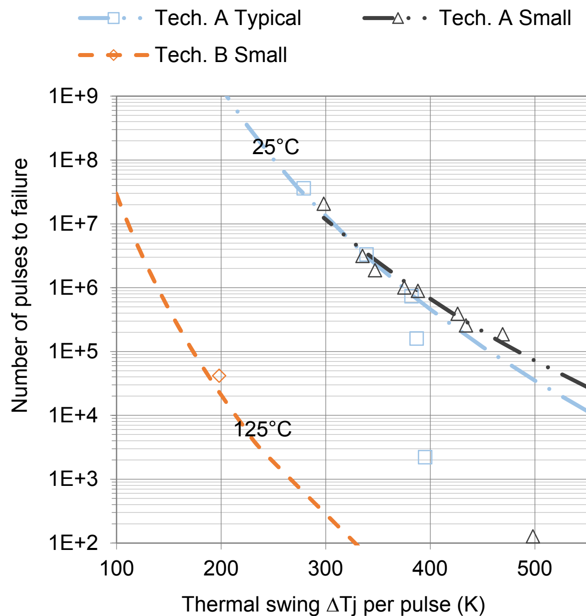

Figure 9: Coffin-Manson chart of mean lifetime (t63 value) from measured test series for different active area of DUT plotted against the thermal swing based on ETHAN simulations. For both device groups the Coffin-Manson model is included and the chuck temperature is indicated.

6. Conclusions

A novel approach for wafer-level test and monitoring of integrity of the multilayer metal-stack in IC process technologies based on the low-cycle fatigue of a power device metallization structure has been described. Employing repetitive power pulsing on wafer-level at the limit of the electro-thermal SOA of the devices reveals changes of intrinsic robustness (e.g. due to process changes, interface weaknesses, or due to layout changes) and is able to activate latent defects. As was shown based on careful exploration of the intrinsic lifetime limits as reference and backed-up by electro-thermal simulation the method may be applied directly on integrated IC-product power stages allowing monitoring on wafer with large area coverage. Additionally portability to small test vehicles which may be integrated in process control module has been shown which allows usage of the method in case respective IC product does not contain suitable devices.

Feasibility of the method has been investigated exemplarily for two smart-power technologies with mixed AlCu and Cu backend stack. Here level of intrinsic robustness decides if the test may be implemented as an end-of-life test with feedback on systematic trends and homogeneity, as we demonstrated by a case-study, or if it may be implemented as a pass/fail test which nevertheless will flag impaired integrity of backend stack. For state-of-the-art smart-power technologies with modular backend options we expect the latter to apply in particular if thin AlCu levels or dual-damascene-copper levels are employed [4].

The methods described are useful in particular to support development and quality for IC products suspected to RPP stress in application, like integrated valve drivers for automotive injection and breaking systems [7]. In this case the additional challenging step is to establish link of low-cycle failure behavior (typically covered sufficiently by Coffin-Manson like modelling) to RPP robustness at use-conditions. But in general our approach allows test vs. reference and monitoring activity of IC process backend-of-line independently from whether or not the respective products of a given process will actually operate under RPP stress conditions in application.

Acknowledgments

Part of this work has been conducted in the project eRamp (grant agreement N°621270), co-funded by grants from Austria, Germany, Slovakia and the ENIAC Joint Undertaking. We thank our colleagues Hans-Helmut Kuge, Christian Maier and Franz Dietz for helpful discussions.

References

1. J. M. Bosc, I. Percheron-Garcon, E. Huynh, P. Lance, I. Pages, J. M. Dorkel and G. Sarrabayrouse. *Reliability characterization of LDMOS transistors submitted to multiple energy discharges, Power Semiconductor Devices and IC´s (ISPSD), International Symposium on*, 165-168, 2000.
2. H.V. Nguyen, C. Salm, J. Vroemen, J. Voets, B. Krabbenborg, J. Bisschop, A.J. Mouthaan, F.G. Kuper. *Fast temperature cycling and electromigration induced thin film cracking in multilevel interconnection: experiments and modeling, Microelectronics Reliability*, 42, 1415-1420, 2002.
3. T. Smorodin, J. Wilde, P. Alpern and M. Stecher. *A temperature-gradient-induced failure mechanism in metallization under fast thermal cycling. Device and Materials Reliability, IEEE Transactions on*, 8(3):590-599, Sept 2008.
4. F. Pozzobon, D. Paci, G. Pizzo, A. Buri, S. Morin, F. Carace, A. Andreini, D. Gastaldi, E. Bertarelli, R. Lucchini and P. Vena. *Reliability characterization and FEM modeling of power devices under repetitive power pulsing, Reliability Physics Symposium (IRPS), IEEE International Conference on*, 5C.4.1-5C.4.8, April 2013.
5. F. Meier, C. Schwarz and E. Werner. *Crystal-plasticity based thermo-mechanical modeling of Al-components in integrated circuits, Computational Materials Scienc*, 94, 122-131, 2014.
6. D. Martineaua, C. Levadea, M. Legrosa, P. Dupuyb, T. Mazeaudb, *Universal mechanisms of Al metallization ageing in power MOSFET devices, Microelectronics Reliability*, 54, 2432 – 2439, 2014.

7. W. Kanert. *Reliability Challenges for Power Devices under Active Cycling, Reliability Physics Symposium (IRPS), IEEE International Conference on*, 409-415, April 2009.
8. P. L. Hower, S. Pendharkar, *Short and long-term safe operating area considerations in LDMOS transistors, Reliability Physics Symposium (IRPS), IEEE International Conference on*, 545-550, April 2005.
9. C. Kendrick, R. Stout and M. Cook. *Reliability of NLDMOS transistors subjected to repetitive power pulses, Reliability Physics Symposium (IRPS), IEEE International Conference on*, 651-652, April 2008.
10. M. Pfost, D. Costachescu, A. Podgaynaya, M. Stecher, S. Bychikhin, D. Pogany and E. Gornik. *Small embedded sensors for accurate temperature measurements in DMOS power transistors, Microelectronic Test Structures (ICMTS), IEEE International Conference on*, 3-7, March 2010.
11. M. Ershov, A. Tcherniaev, Y. Feinberg, P. Lindorfer, W. French and P. Hopper. *Numerical simulation of metal interconnects of power semiconductor devices, Power Semiconductor Devices and IC´s (ISPSD), International Symposium on*, 185-188, June 2010.
12. M. Pfost, C. Boianceanu, H. Lohmeyer and M. Stecher. *Electrothermal simulation of self-heating in DMOS transistors up to thermal runaway, IEEE Transactions on Electron Devices*, 60(2), 699-707, Feb 2013.
13. K.C. Norris and A.H. Landzberg. *Reliability of controlled collapse interconnections, IBM Journal of Research and Development*, 13(3):266-271, May 1969.
14. T. Smorodin, J. Wilde, P. Nelle, E. Lilleodden and M. Stecher. *Modeling of DMOS subjected to fast temperature cycle stress and improvement by a novel metallization concept, Reliability Physics Symposium (IRPS), IEEE International Conference on*, 689-690, April 2008.
15. A. Ferrara, J. Claes, M. Swanenberg, L. van Dijk and P. G. Steeneken. *Accelerated resistance degradation in aluminum by pulsed power cycling, Power Semiconductor Devices and IC´s (ISPSD), International Symposium on*, 301-304, May 2015.
16. D. Simon, C. Boianceanu, G. De Mey and V. Topa. *Experimental reliability improvement of power devices operated under fast thermal cycling, Electron Device Letters, IEEE*, 36(7):696-698, July 2015.
17. W. Muth and W. Walter. *Bias temperature instability assessment of n- and p-channel MOS transistors using a polysilicon resistive heated scribe lane test structure, Microelectronics Reliability*, 44, 1251-1262, 2004.
18. G. Pobegen, M. Nelhiebel, S. de Filippis, and T. Grasser. *Accurate high temperature measurements using local polysilicon heater structures, Device and Materials Reliability, IEEE Transaction on*, 14(1), 169, March 2014.