A Temperature-Gradient-Induced Failure Mechanism in Metallization Under Fast Thermal Cycling

Tobias Smorodin, Jürgen Wilde, Peter Alpern, and Matthias Stecher

***Abstract*—In this paper, a novel mechanism is shown to cause the failure evolution in a metallization system under fast temperature cycle stress. The failure evolution is triggered by the lateral temperature distribution across the device, which causes an accumulating plastic deformation of the metallization. The root cause for the deformation emerges at the position of the maximum gradient in temperature.**

***Index Terms*—Interlayer dielectric (ILD) cracking, metallization failure, power cycling, short circuit, temperature cycling.**

I. Introduction

A BROAD range of applications in the automotive area requires the switching of inductive loads by electronic control units. With engine management, for instance, the injection valves are electromagnetically operated, whereas the inductive loads are driven by smart-multichannel *double-diffused MOS* (DMOS) switches. A DMOS transistor that is broadly used in this field is the quasi-vertical DMOS transistor, which was introduced with the first BCD technology [1].

With the application named earlier, the DMOS switches an inductive load, which opposes the current change by a rise of its terminal voltage (Fig. 1). Therefore, the Zener clamping between drain and gate prevents the DMOS from avalanche, and the current can decay at the adjusted clamping voltage. The drain–source voltage (V_{DS}) in the OFF state is around 100 times higher than in the ON state. Therefore, although the current decays, a high initial power dissipation causes a sharp temperature rise within a fraction of milliseconds. The temperature range considered in this paper is below the single pulse destruction limit reported in [2].

The failure mechanism considered in this paper evolves in the metallization of the DMOS, during a high number $(10^4$–$10^9)$ of repetitive temperature pulses [3]. As the metallization of the DMOS has a higher coefficient of thermal expansion than the underlying silicon substrate, a thermomechanical stress is introduced during fast temperature cycling. This causes a viscoplastic deformation of the aluminum, which accumulates with each switching cycle. After a high number of switching cycles, the deformation leads to cracking of the passivation and the interlayer dielectric (ILD). The consequence is extruding aluminum that causes electric short circuits and the device failure (Fig. 2). With a previous publication, we have shown that electromigration has no impact on the time to failure of the DMOS [4].

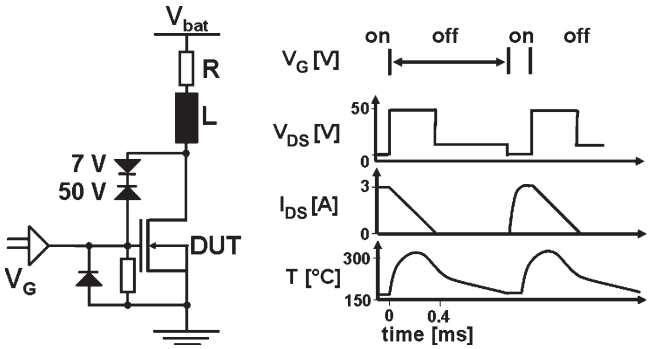

Fig. 1. Circuit diagram of a low-side switch driving an inductive load. The relationship between V_{DS}, I_{DS}, and the temperature is shown schematically.

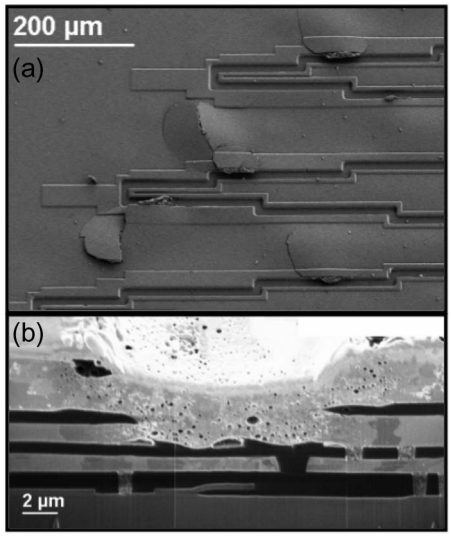

Fig. 2. (a) SEM image of a DMOS switch with aluminum in the power metallization after exposure to fast temperature cycle stress. Buckling of the metallization and passivation cracks are visible from which aluminum is extruding. (b) FIB cross section of a failure spot: A power overshoot results from the short circuit and leads to a local melting and evaporation of the aluminum.

As the processing cost is increasing for advanced Smart Power Technologies, there is an increasing demand for further shrinkage of the devices. One approach is the lateral DMOS, which has a different layout and the advantage of a significantly lower specific on-resistance $R_{\mathrm{DS(on)}} \cdot \mathrm{A}$ [5]. For the aspired applications, however, the size of the devices is not only limited by the $R_{\mathrm{DS(on)}} \cdot \mathrm{A}$ but also by the thermal boundary conditions.

Manuscript received November 21, 2007; revised March 4, 2008. First published August 12, 2008; current version published October 16, 2008.

T. Smorodin, P. Alpern, and M. Stecher are with Infineon Technologies AG, 85579 Neubiberg, Germany (e-mail: tobias.smorodin@infineon.com; peter.alpern@infineon.com; matthias.stecher@infineon.com).

J. Wilde is with the Department of Microsystems Engineering, University of Freiburg, 79110 Freiburg, Germany (e-mail: wilde@imtek.de).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TDMR.2008.2002359

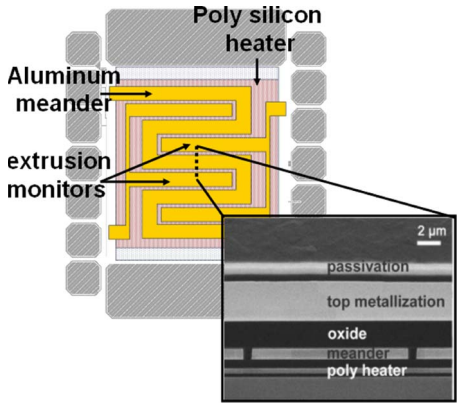

Fig. 3. Schematic view of the test structure: With the polysilicon heating plate, the fast temperature cycling (10–100 Hz) of the metallization is possible without current flow in the metallization. The inset shows an FIB cross section of the test structure.

Therefore, it is of great interest to increase the thermal capabilities of power DMOS stages under fast (10–100 Hz) temperature cycle stress. The key to an improvement is a fundamental understanding of the failure mechanism, which is presented in this paper.

II. Experimental

The test structure used for this study consists of a quadratic plate of polysilicon, which is electrically contacted from two sides. By applying a pulsed voltage on that resistor, it can be used as a miniaturized heating plate. Above the plate, a meandering aluminum line is located. This meander is traversed from both sides by extrusion monitors, which allow the detection of lateral short circuits (Fig. 3) [6]. The meander and the extrusion monitors have the same linewidth and are separated by 0.8-μm wide oxide walls. To evaluate the influence of the conductor path design on the failure behavior, structures with linewidths of 3.2 and 12.8 μm, respectively, are investigated. The meandering line and the extrusion monitors are covered by a stack of silicon oxide and aluminum, which allows the detection of vertical short circuits (inset Fig. 3).

As the temperature within the metallization system is decisive for the stress level, it is measured with an additional test structure. It consists of a very narrow aluminum meander ($R \sim 900\ \Omega$ at room temperature), located directly above the center of the polysilicon plate. In Fig. 4, the temperature response to a 5-ms power pulse is shown. The drift of the sensor room temperature resistance does not exceed 1% after 1 hr of measurement.

With this test chip, it is possible to temperature cycle the metallization quickly and without any current flow in the metallization. Lateral and vertical short circuits are taken as failure criteria. After the detection of the first short circuit, the fast temperature cycling is stopped immediately. This assures that the failure situation is preserved for further investigation.

III. Results

The loading condition is $\Delta T = 300$ K, where ΔT is the maximum temperature rise during one power cycle. The package temperature reaches a steady-state value of $T = 125\ ^{\circ}$C after approximately 30 min. A Weibull plot of the time-to-failure data for the two conductor line sizes is shown in Fig. 5. The Weibull distribution is applicable when the weakest of many flaws propagates to failure. This is appropriate to fast temperature cycling because the device failure is defined as the first electric contact between two conductor lines or between a conductor line and the covering plate.

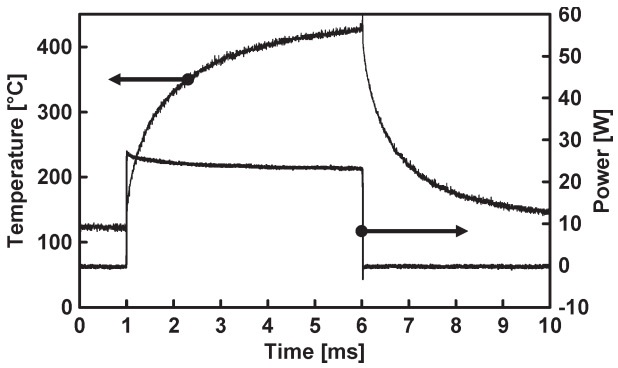

Fig. 4. Power pulse with a duration of 5 ms applied to the polysilicon heater and resulting temperature profile measured in the metallization.

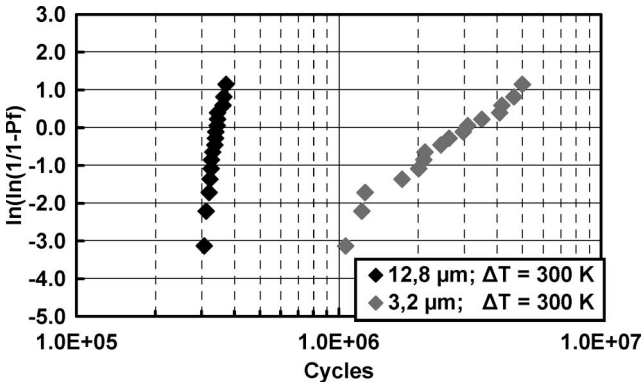

Fig. 5. Weibull plot of time-to-failure data for two different line sizes.

A variation in mean time to failure is observed for the different line sizes. The 3.2-μm structures show an increased time to failure in comparison with the structures with a linewidth of 12.8 μm. Another significant difference is observed for the slope of the failure data. For the 3.2-μm structures, the failure distribution is broader in general ($\beta = 2$).

A. DICM Analysis

The first analysis step of the failed devices was done by differential interference contrast microscopy (DICM), which provides access to the change in the topography of the devices. To evaluate the deformation behavior of the meander layer, the passivation and the covering aluminum plate were removed. Likewise, the position of the short circuit can be localized. For the 12.8-μm structures, all short circuits are located within the circle shown in Fig. 6. The DICM reveals strong buckling of the conductor lines, whereas the deformation is maximal in the center of the structure and declines toward the edges.

The same analysis was performed for the devices with a linewidth of 3.2 μm. In this case, the failure spots are not

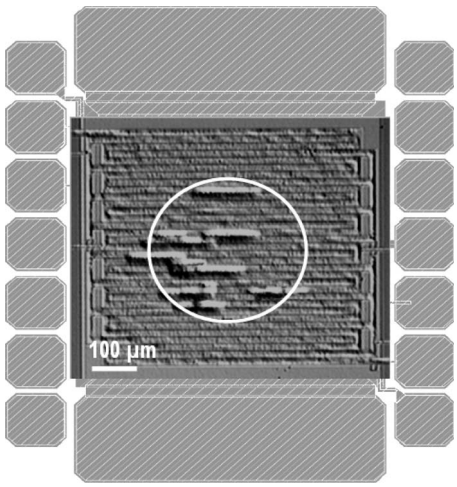

Fig. 6. DICM image of a stressed device with a linewidth of 12.8 μm. The failure mapping for 12 devices reveals a concentration of the short circuits within the displayed circle.

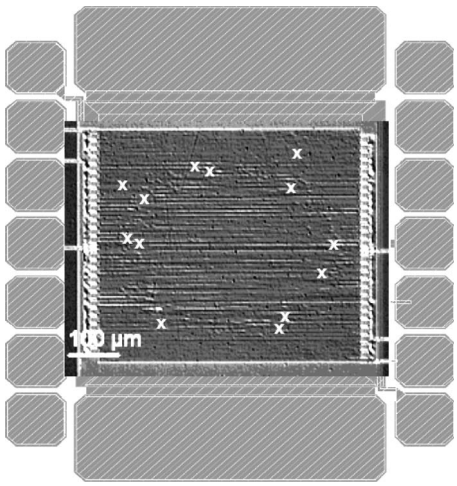

Fig. 7. DICM image of a stressed device with a linewidth of 3.2 μm. The failure mapping (see crosses) shows an alignment of the failure spots circularly around the center.

located in the center of the structure but circular around the center. The DICM shows no buckling of the lines in the center but a ring-shaped deformation on a circle around the center where the devices finally fail (Fig. 7).

The failure spot is not necessarily located where the maximum temperature rise occurs during the pulsed heating. The temperature distribution is nonuniform across the plate, and the highest temperature is observed in the center. In the case of the 3.2-μm-wide lines, the failure spots are not located in the center. Instead, they are found where the maximum plastic deformation of the lines is observed in the DICM image. In the case of the 3.2-μm-wide lines, the maximum plastic deformation is located on a circle around the center (Fig. 7).

B. FIB Analysis

The FIB cross sections of the failure spots are prepared for a further investigation of the failure evolution. A cross-sectional view of a device with 12.8-μm linewidth is shown in Fig. 8. It reveals a lateral crack, which propagates from the

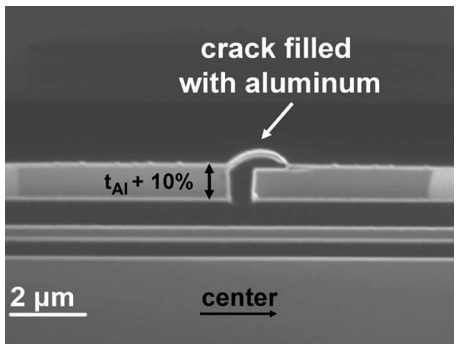

Fig. 8. FIB cross section of a short circuit between two neighboring 12.8-μm-wide conductor lines. Thickness measurement shows an increase in line thickness of up to 10%.

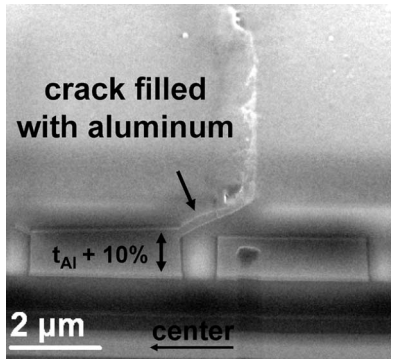

Fig. 9. FIB cross section of a short circuit starting at the edge of a 3.2-μm-wide conductor line. Thickness measurement shows an increase in line thickness of up to 10%.

edge of one conductor line to the neighboring line. The crack points are toward the center of the test structure. A distance measurement is performed on the images of the cross sections to determine the height of the conductor lines. It shows an increase in aluminum thickness of up to $\sim$10% in comparison to the unstressed specimens. This confirms the observation of the DICM. The crack is filled with aluminum, which was verified via EDX.

The same preparation was done for a failed device with a linewidth of 3.2 μm. The result is shown in Fig. 9. In this case, the crack is oriented toward the side of the test structure. Again, a thickness measurement reveals an increase in thickness of the aluminum line of up to $\sim$10% at the location of the short circuit.

These results indicate that the plastic deformation of the conductor lines plays a central role for the failure evolution, as the failure location coincides with a significant increase in thickness of the lines. The mechanism of the plastic deformation under fast temperature cycle stress is therefore investigated on the level of the covering aluminum plate. As differential images of the DICM images are taken, the contrast is even further enhanced.

This reveals that, within the first ten temperature cycles, a ring-shaped elevation forms in the covering aluminum at a certain distance to the center (Fig. 10). The position of this inhomogeneity in thickness shows a correlation to the temperature distribution across the test structure. The temperature profile across the aluminum plate is displayed to demonstrate this. It was obtained with a thermal infrared camera during

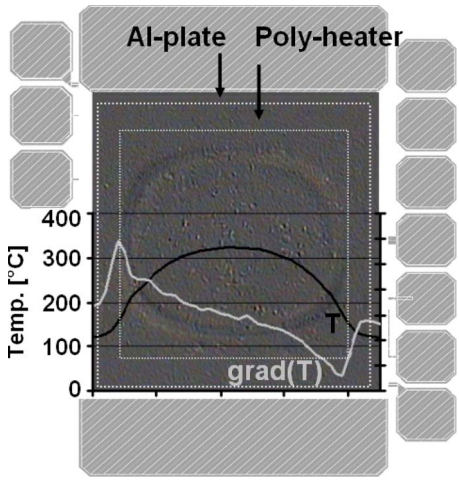

Fig. 10. Ring-shaped elevation is initiated in the covering aluminum layer during the first ten temperature cycles ($\Delta T = 300$ K, $t_{\mathrm{pulse}} = 5$ ms, and $f = 10$ Hz). It becomes visible after the subtraction of the initial state DICM image. A temperature profile through the center of the test structure is displayed additionally.

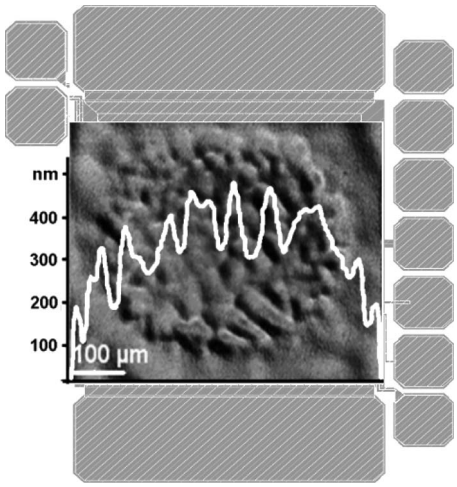

Fig. 11. DICM image of the covering aluminum layer after 10^5 loading cycles. The height profile through the center of the structure was obtained with a mechanical profilometer.

pulsed heating (*Thermovision 900*, line frequency: 2.5 kHz). The temperature is highest in the center of the polyheater and declining toward the edges. At the edge of the polyheater, the gradient in temperature becomes maximal. The size of the covering aluminum plate exceeds the size of the polyheater (see dimensions in Fig. 10). The plastic deformation is initiated in the aluminum layer, exactly where the gradient in temperature is maximal. With an increasing number of loading cycles, the deformation is continuously shifted toward the center of the structure. In Fig. 11, it can be seen that, after 10^5 loading cycles, the layer thickness in the center increases by up to 6% ($t_{\mathrm{Al}} = 3.5$ μm). Furthermore, inhomogeneities in layer thickness evolved with a height of up to 150 nm. In this state, the ring-shaped substructure is still visible. This results from the lateral temperature distribution.

The same deformation mechanism is underlying for the meandering layer, but it evolves differently, depending on the line size. For the 12.8-μm-wide lines, the maximum deformation is found in the center, as observed with the covering plate. For the 3.2-μm-wide lines, a much higher amount of oxide bars between the lines (0.8-μm oxide for 3.2-μm aluminum => 20% oxide versus 6% for the 12.8-μm-wide lines) prevents the propagation toward the center. Therefore, the maximum deformation is located circularly around the center. The linewidth dependence of the failure behavior was discussed in detail in [7]. In the next step, the deformation behavior is analyzed with a thermomechanical simulation.

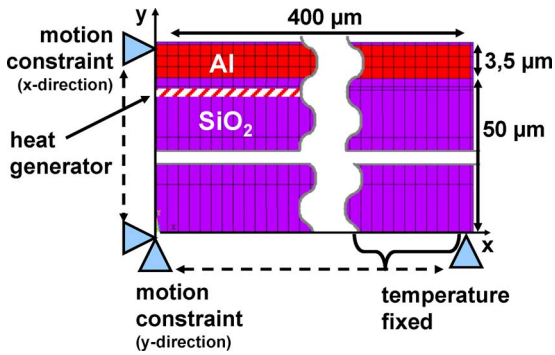

Fig. 12. Plane FEM model of an aluminum layer on silicon oxide substrate. With the thermal simulation, a temperature profile is generated, which is then applied to the subsequent mechanical simulation.

IV. Simulation

A. Bimaterial Model

The impact of the lateral temperature distribution on the deformation behavior is studied with a sequential thermomechanical analysis using the finite element code *ANSYS*. With the first set of simulations, only the top metallization plate is considered as the simple geometry that enables the evaluation of the underlying mechanism. Later, this model is extended, such that the conductor lines (meander layer) are considered also.

With the first step, a nonuniform temperature distribution is generated with a thermal simulation (element type: *plane55*; real constants: *none*; element behavior: *plane*). The resulting temperature field is then applied to a mechanical analysis (element type: *plane182*; real constants: *none*, element behavior: *plane strain*). A 2-D model is generated in order to be able to simulate an increased number of temperature pulses.

The plane-strain model consists of a 400-μm-wide oxide bar with a thickness of 50 μm. The oxide bar is covered by a layer of 3.5-μm-thick aluminum. To mimic the metallization system, the aluminum is covered by a 0.3-μm-thick oxide layer. For the thermal simulation, power dissipation is applied to a layer of elements, which are located within the silicon oxide. The temperature of the nodes at the right side of the model is fixed to 125 °C in order to generate a gradient in temperature in the x-direction (Fig. 12).

With the subsequent mechanical simulation, the model is cycled between a uniform temperature of 125 °C (= OFF state) and the temperature distribution obtained with the thermal simulation (= ON state). For the experiments shown in the previous section, dies with a thickness of 380 μm were used. Due to the stiffness of the silicon die, a significant motion of the die/metallization interface in the y-direction is suppressed.

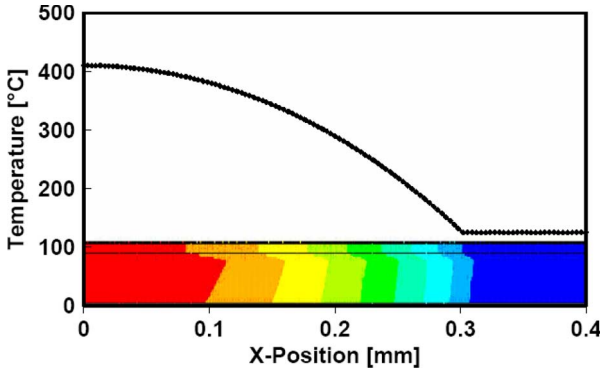

Fig. 13. Temperature distribution across the bimaterial model at the end of a 5-ms power pulse.

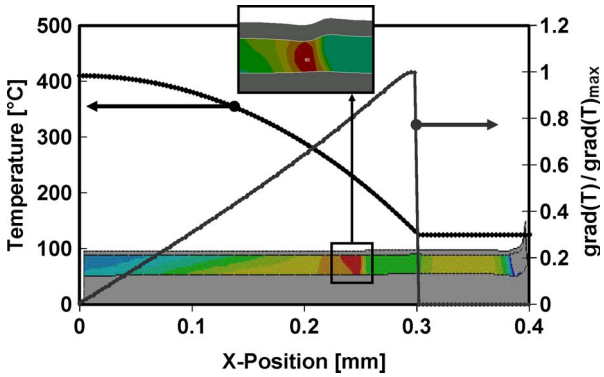

Fig. 14. Equivalent plastic strain within the aluminum layer in the OFF state ($T = 125$ °C) after ten preliminary loading cycles. Close to the location of the maximum gradient of temperature, the maximum amount of plastic deformation is observed.

Therefore, the displacement of the nodes at the lower border of the model is constrained in the y-direction in order to provide a stiff frame for the model.

The material properties of the silicon oxide are assumed to be linear–elastic and temperature independent. A temperature-dependent nonlinear kinematic hardening model (Chaboche) is applied to model the plasticity of the aluminum [8], [9].

B. Results

A plot of the temperature distribution at the end of the 5-ms power pulse is shown in Fig. 13. The highest temperature is found in the center of the model. It is declining to the side that results in a considerable gradient of temperature in the x-direction.

Fig. 14 shows a plot of the equivalent plastic strain in the OFF state ($T = 125$ °C; homogenous) after ten loading cycles. The maximum deformation is located in the aluminum layer that is close to the location of the maximum gradient in temperature during the ON state. For comparison, the temperature distribution and the gradient of temperature are displayed additionally. The inset in Fig. 14 shows the deformation of the layer in detail. This location of the deformation at the maximum gradient in temperature coincides well with the experimental observations shown in Section III (Fig. 10).

When the model is further cycled between the OFF state ($T = 125$ °C; homogenous) and ON state, the deformation is shifted toward the center of the structure. This results in an increasing deflection of the nodes in the y-direction. As a result, the thickness of the aluminum layer is increasing in the center and decreasing toward the edges. A plot of the resulting vertical shift of the surface is shown in Fig. 15 for increasing the number of loading cycles.

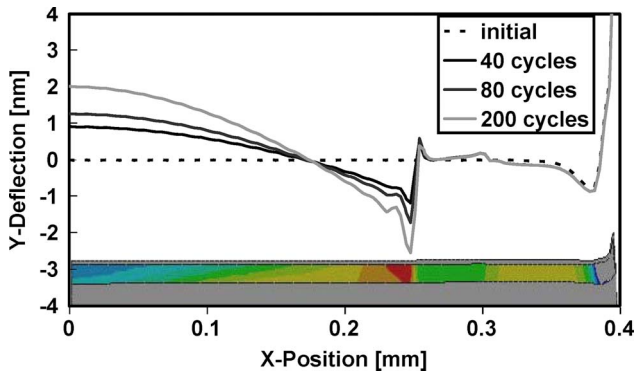

Fig. 15. Vertical (y) deflection of the nodes as a function of the position along the surface of the aluminum in the cooled state ($T = 125$ °C). For comparison, the equivalent plastic strain after ten preliminary loading cycles is shown additionally.

V. Discussion

In the previous section, the temperature-gradient-driven plastic deformation of an aluminum layer on an elastic substrate was analyzed with FEM simulation. During the first few temperature cycles, a plastic deformation is introduced in the film, where the maximum gradient of temperature emerges. This is the starting point for the subsequent deformation of the film. The amount of accumulated plastic deformation, as well as the deflection of the surface in the center, overcomes this initial distortion with an increasing number of cycles.

The root cause for the initiation of the plastic deformation at the position of the maximum gradient of temperature is obtained from the evaluation of the shear stress within the aluminum layer. In Fig. 16, a change in the sign of the shear stress is observed at the location of the maximum gradient in temperature. The direction of the shear stress is changing from toward the center within the high temperature region to outward in the low temperature region. The absolute value of the shear stress changes with the temperature during one temperature cycle, but the shift in direction remains unchanged. At the border of the model ($x = 0.4$ mm), the shear stress is increasing due to an edge effect, which causes a deflection of the surface (see Fig. 15). However, as the temperature is fixed in this region, the surface deflection at the edge is not altered by the thermal cycling.

A ratcheting deformation of thin metal films under cyclic loading has been described by Huang *et al.* for packaged devices that are exposed to passive temperature cycling [10]. The premise for ratcheting deformation is a combination of two stress components, whereas the cycling component is superimposed by a component that points to a fixed direction. In the case of the passive temperature cycling, a shear stress points toward the center of the chip, which is a residual of the packaging process. Although this shear stress is below the

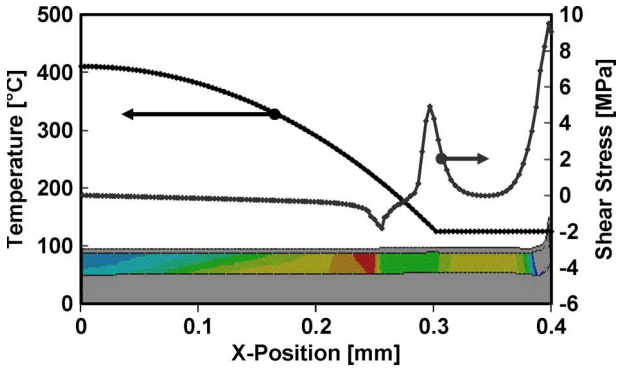

Fig. 16. Temperature profile along the bimaterial beam and the resulting shear stress. For comparison, the equivalent plastic strain after ten preliminary loading cycles is shown additionally.

yield stress of the metal film, one-directional shifting of the film is observed. The reason is that the metal is yielding at low temperatures due to the different thermal expansions in comparison to the substrate. The small shear stress then shifts the film incrementally with each cycle such that a high amount of accumulated plastic deformation results after an increased number of cycles.

The test structures of this publication are assembled in ceramic packages without molding compound. Hence, no residual shear stresses act on the metallization. With the experiments considered in this paper, the one-directional shear stress is no residual from the assembly but an intrinsic effect which results from the lateral temperature distribution. The deformation behavior shows no correlation to the position of the metallization with respect to the coordinates of the chip but to the temperature distribution within the test structure. The orientation of the shear stress during a temperature pulse is shown schematically in Fig. 17. The resulting shear stress is in the range of up to 5 MPa and is therewith significantly below the yield stress of the metallization ($\sigma_{\mathrm{yield}} \sim 120$ MPa at room temperature). Nevertheless, the metallization is ratcheting a small amount toward the center of the structure with each cycle. The reason is that the metal film yields due to the different thermal expansions in comparison to the substrate. The strain–stress hysteresis in Fig. 18 is evaluated in the center of the structure over 50 loading cycles, whereat each tenth cycle is displayed. Due to the asymmetrical temperature profile (see Fig. 4), a different shape is obtained for heating and cooling. The graph shows the cyclical yielding of the metallization. Furthermore, it illustrates the translation of the strain–stress hysteresis and, therewith, the accumulation of plastic deformation.

As the mechanical yield stress of aluminum is decreasing with increasing temperature, the yield stress of the metallization is lowest in the center. Therefore, the ratcheting deformation of the film evolves faster in the center of the structure, which is seen experimentally (Fig. 11) and with the thermomechanical simulation (Fig. 15). Without assuming a temperature-dependent material model for the aluminum, the deformation is also initiated at the location of the maximum gradient, but the film thickness in the center stays nearly unchanged with cycling (Fig. 19).

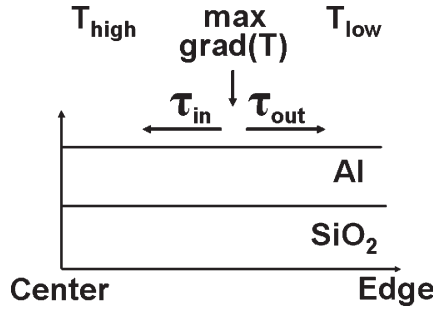

Fig. 17. Schematic view of the shear stress distribution resulting from the nonuniform temperature profile during a temperature pulse.

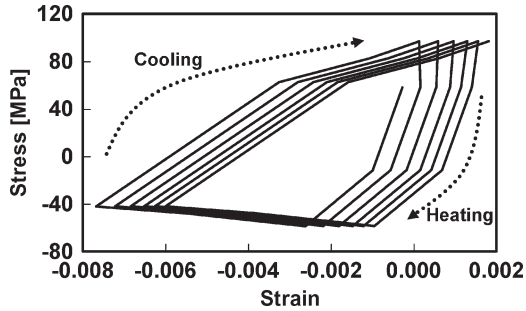

Fig. 18. Strain stress hysteresis for the aluminum in the center of the structure during 50 loading cycles (each tenth cycle is displayed).

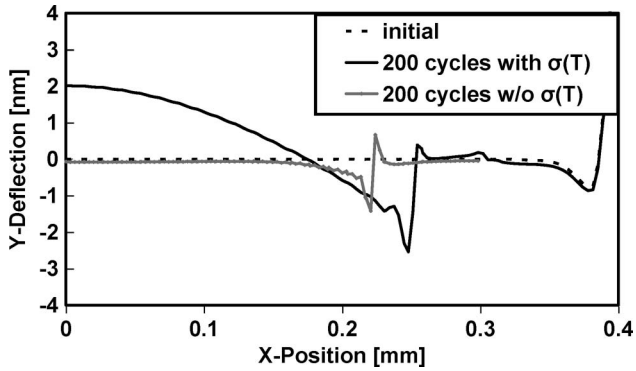

Fig. 19. Vertical (y) deflection of the nodes as a function of the position along the surface of the aluminum at homogenous temperature $T = 125$ °C. The simulation is performed with and without assuming a temperature-dependent yield stress for the aluminum.

A. Buckling of the Aluminum

With the experimental results shown in Section III, the evolution of the surface deflection is observed with a mechanical profilometer. In Fig. 11, a DICM image and the corresponding surface profile are displayed. This shows the bulging of the surface in the center of the structure, as also seen with the thermomechanical simulation. Furthermore, an additional modulation of the surface evolves, which can be described as the buckling of the surface. In Fig. 20, an atomic force microscope (AFM) image of those surface buckles is shown.

The AFM image confirms the observation of the mechanical profilometer; buckles with a height of up to 150 nm are observed, whereas the layer thickness is 3.5 μm. However, with the thermomechanical simulation, these local variations in thickness are not observed. The deflection of the film is

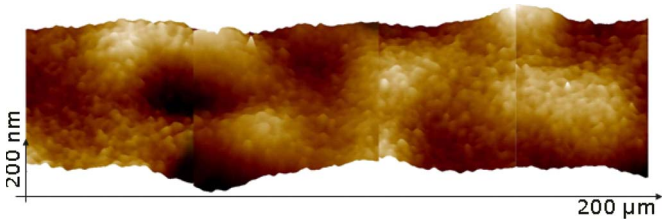

Fig. 20. Three-dimensional plot of a scanning profile obtained with AFM shows a significant buckling of the surface.

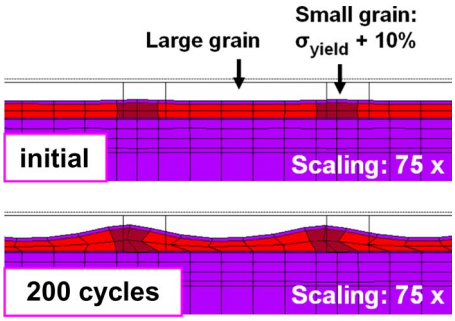

Fig. 21. Model of a metallization layer with a varying microstructure. For the sections that exhibit small grains, a yield stress, which is increased by 10%, is assumed. After 200 cycles, a buckling of the surface evolves, whereas the larger grains are more deformed.

positive in the center of the structure and negative at the side as the material is shifted toward the center. The resulting surface, however, is smooth compared with the experimental finding.

The main difference between the experiment and the simulation is that a perfectly homogenous aluminum layer is assumed for the simulation. In reality, the aluminum layer is composed of grains with different sizes that have also varying mechanical properties. According to the Hall–Petch relationship, the change in mechanical yield stress σ_{yield} is proportional to the inverse square root of the grain size d.

In order to evaluate whether local variations of the yield stress across the layer can cause buckling, an additional simulation is performed. The continuous aluminum layer shown in Fig. 12 is split up into two types of alternating segments with different widths and yield stresses. The model is shown in Fig. 21.

The segments have widths of 15 and 4 μm, and the yield stress is increased by 10% for the small grains. The model is loaded with the same thermal conditions that are applied for the mechanical simulation in Section IV. A plot of the resulting evolution of the surface deflection is shown in Fig. 21.

The alternating yield stress along the segments is the starting point for the evolution of strong buckling with increasing amplitude. Fig. 22 shows how the interaction between "soft" and "hard" grains contributes to the evolution of the inhomogeneities. The "soft" grains become more compressed and form the minima, whereas the "hard" grains evolve to maxima. The simulation shows that the experimentally observed buckling could be caused by the different contributions of individual grains to the overall deformation. The alternating yield stress adds an additional modulation in thickness to the deflection shown in Section IV.

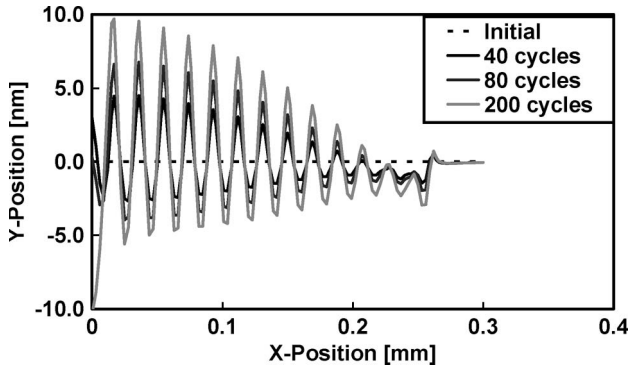

Fig. 22. Deflection of the nodes along the surface of the aluminum at homogenous temperature $T = 125$ °C. With the assumption of a varying microstructure (see Fig. 20), an additional buckling is superimposing the deflection of the surface.

Another reason for the experimentally observed surface modulation could be the wrinkling of an elastic film on a metal layer due a ratcheting deformation [11]. Im and Huang have shown that a compressively strained film, which is bonded to a metal layer, induces normal and shear tractions at the film–metal interface. As a consequence, the wrinkling of the film grows during temperature cycling, as plastic deformation is accumulated in these preset directions. With the simulations shown in Section IV, the presence of passivation alone was not observed to cause a buckling (Fig. 15). One possible reason for this could be an insufficient modeling of the film prestress.

B. Deformation of the Conductor Lines

In Section III, the plastic deformation of the aluminum lines is shown to be the root cause for the device failure. The bimaterial model is therefore extended to investigate the failure evolution within the test structures. Therefore, an additional layer is added, which represents the meander and the extrusion monitors (Fig. 23). This 2-D model of the test structures with 12.8-μm-wide lines is cycled between a homogenous temperature (OFF state) and a nonuniform temperature distribution (ON state), as shown with the bimaterial model in Section IV.

The cycling leads to an accumulation of plastic deformation within the covering aluminum layer, as well as within the conductor lines [Fig. 24(a)]. In Section III, it is shown that the failure of the devices evolves within the meandering layer. The devices fail due to vertical or lateral short circuits, which start in both cases at the edges of the conductor lines and propagate through the ILD (Figs. 8 and 9). In Fig. 24(a), it can be seen that the plastic deformation within the meandering layer is maximum for the lines in the center. Therefore, the accumulated plastic deformation is evaluated for the line marked in Fig. 24(a). The plastic deformation is continuously increasing with increasing number of loading cycles [Fig. 24(b)].

In order to evaluate the effect of the plastic deformation on the stress level in the ILD, the first principal stress is also analyzed from this simulation [Fig. 24(c)].

The stress level within the ILD is continuously increasing with the inceasing number of temperature cycles, and from approximately 1000 cycles on tensile only, loading is observed. The amplitude between OFF and ON states is almost constant,

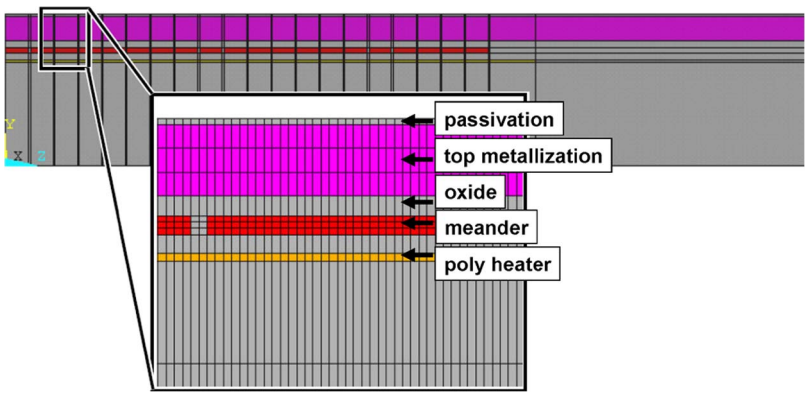

Fig. 23. Plane FEM model of the test structure including the covering metallization plate and 12.8-μm-wide conductor lines.

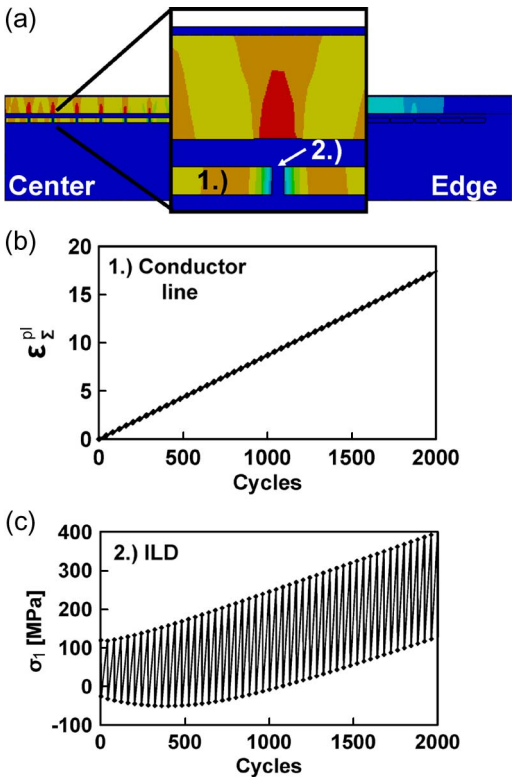

Fig. 24. (a) Equivalent plastic strain within the metallization of the test structure after the simulation of 2000 cycles. (b) Evolution of the accumulated plastic deformation within the cross section of one conductor line during 2000 loading cycles. (c) Resulting stress within the ILD is continuously increasing with the number of cycles.

and the maximum stress value is continuously increasing. The increasing stress level finally leads to the observed cracking of the ILD. Due to high pressure within the conductor line, the crack is immediately filled with aluminum, and the device fails due to an electric short circuit (Figs. 8 and 9).

VI. Summary and Failure Model

The experimental investigations in Section III have shown that the device failure does not necessarily occur where the maximum temperature rise is observed. Instead, the failure spots are located where the maximum plastic deformation is observed with DICM (Figs. 6 and 7). The FIB images of these failure spots reveal a crack in the ILD, which is filled with aluminum and accounts for the short circuit (Figs. 8 and 9). The thickness measurement of the aluminum lines at the failure spot shows an increase in thickness of up to 10%. This coincides with the observation from the DICM. The viscoplastic deformation of the aluminum lines is identified to be the root cause for the failure evolution. Furthermore, it was shown that the nonreversible deformation is caused by the temperature distribution across the test structure. The deformation is initiated where the maximum gradient in temperature is observed in plane. The FEM analysis of this deformation mechanism shows that a change in shear direction occurs at the position of the maximum gradient in temperature. This shear stress is biasing the temperature cycling and causes a continuous displacement of the aluminum toward the high temperature region.

With the thermomechanical analysis of the test structure in Section IV, it is seen that the viscoplastic deformation of the aluminum lines causes an increasing stress within the ILD. All failure analysis methods including FIB and TEM images did not reveal any evolution of submicrocracks at the edges of the aluminum lines. Only one catastrophic crack filled with aluminum is found, which accounts for the short circuit. This indicates that the ILD cracking is not caused by a cyclic fatigue of the ILD itself, which would lead to a continuous evolution of microcracks. Instead, overstress, caused by the locally increasing aluminum thickness, can be identified as the source of defect.

Based on these findings, the following failure hypothesis is suggested.

The initial state of the samples is represented by plane conductor lines, which are surrounded by silicon oxide [Fig. 25(a)]. The fast temperature cycle stress leads to the viscoplastic deformation of the conductor lines, which is driven by the lateral gradient in temperature. As a result of this deformation, inhomogeneities in the thickness of the aluminum lines evolve [Fig. 25(b)], whereat the location of the inhomogeneities is strongly influenced by the layout of the conductor lines. Additional fluctuations in layer thickness can result from the variations of the aluminum microstructure, e.g., grain size. A rising mechanical stress within the ILD results particularly at those locations where the thickness of the aluminum lines is increased. An additional stress is caused by the different thermal expansions of aluminum and oxide during each single pulse [Fig. 25(c)]. Beyond a critical stress in the ILD, a crack is

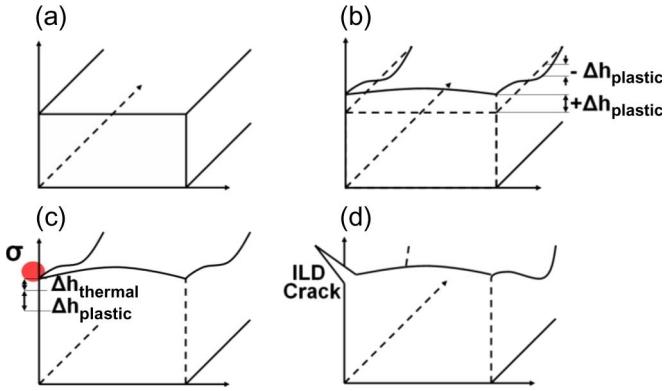

Fig. 25. (a) Schematic view of one conductor line embedded in silicon oxide. (b) Fast temperature cycling causes a plastic deformation of the conductor line that leads to inhomogeneities in the thickness of the conductor line. (c) Locally increasing layer thickness causes an increasing stress—particularly at the edges. In addition, a stress is caused from the thermal strain during the pulses. (d) Beyond a critical stress limit, ILD cracking results. The crack is immediately filled with aluminum that results in a short circuit.

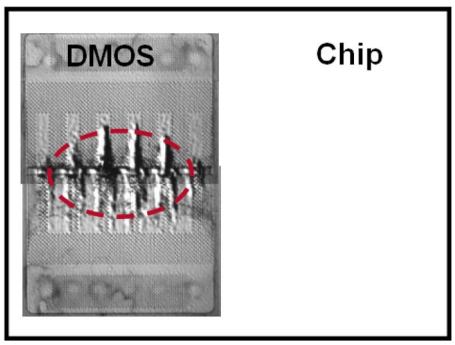

Fig. 26. Deformation behavior observed with the metallization of stressed DMOS is oriented circularly around the center. Moreover, the DMOS was located at the border of a larger chip and packaged in mold compound; the deformation shows no correlation to the position on the chip.

formed [Fig. 25(d)], which is immediately filled with aluminum due to the high hydrostatic pressure within the conductor line. The failure of the device is detected as a short circuit.

VII. Outlook

In this paper, a new deformation-driven failure mechanism of semiconductor dielectrics is introduced. The findings with the thermomechanical test structures can now be transferred to a DMOS switch, which is subjected to repetitive temperature pulses during operation. A DICM image of the metallization for a stressed device is shown in Fig. 26. The DMOS is located at the edge of the chip, but the trait of the deformation points toward the center of the DMOS, which is comparable to the test structures (Fig. 6). Moreover, in the case of the DMOS, the deformation of the metallization is driven by the temperature distribution across the device.

This investigation was motivated by the special application of repetitive inductive load switching with power devices. However, in the future, the increasing packing density and, therefore, increasing power density will lead to severe temperature rises also with consumer electronics. This basic investigation of the deformation mechanism and the failure behavior will therefore be of increasing importance also beyond power devices.

References

[1] B. Murari, "Recent development and trends in intelligent power IC technologies," in *Proc. IEEE Custom Integr. Circuit Conf.*, 1986, pp. 134–137.
[2] M. Denison *et al.*, "Influence of inhomogeneous current distribution on the thermal SOA of integrated DMOS transistors," in *Proc. ISPSD*, 2004, pp. 409–412.
[3] J. M. Bosc *et al.*, "Reliability characterization of LDMOS transistors submitted to multiple energy discharges," in *Proc. ISPSD*, 2000, pp. 165–168.
[4] T. Smorodin *et al.*, "Power-cycling of DMOS-switches triggers thermomechanical failure mechanisms," in *Proc. ESSDERC*, 2007, pp. 139–142.
[5] R. S. Wrathall *et al.*, "Charge controlled 80 V lateral DMOSFET with very low specific on-resistance designed for an integrated power process," *IEDM Tech. Dig.*, pp. 954–957, 1990. Tech. Dig. Cat. No. 90CH2865-4.
[6] H. V. Nguyen *et al.*, "Fast thermal cycling stress and degradation in multilayer interconnects," in *Proc. 9th IPFA*, 2002, pp. 135–139.
[7] T. Smorodin *et al.*, "Investigation and improvement of fast temperature cycle reliability for DMOS-related conductor path design," in *Proc. IEEE IRPS*, 2007, pp. 486–491.
[8] S. Bari *et al.*, "Anatomy of coupled constitutive models for ratcheting simulations," *J. Plast.*, vol. 16, no. 3, pp. 381–409, 2000.
[9] P. J. Armstrong *et al.*, "A mathematical representation of the multiaxial Bauschinger effect," Central Electricity Generating Board, Berkeley, U.K., CEGB Report RD/B/N 731, 1966.
[10] M. Huang *et al.*, "Metal film crawling in interconnect structures caused by cyclic temperatures," *Acta Mater.*, vol. 49, no. 15, pp. 3039–3049, Sep. 2001.
[11] S. H. Im and R. Huang, "Ratcheting-induced wrinkling of an elastic film on a metal layer under cyclic temperatures," *Acta Mater.*, vol. 52, no. 12, pp. 3707–3719, 2004.

Tobias Smorodin studied physics at the Ludwig-Maximilians-Universität Munich, Munich, Germany and the University of Nottingham, Nottingham, U.K. He received the Diploma degree in 2004. He is currently working toward the Ph.D. thesis at the University of Freiburg, Freiburg, Germany.

Since 2005, he has been with Infineon Technologies AG, Neubiberg, Germany. His field of research is reliability physics of power-related metallization systems.

Jürgen Wilde received the Dipl.Ing. degree in materials science from the Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany, and the Ph.D. degree from the Technical University of Clausthal, Clausthal-Zellerfeld, Germany.

He is currently a Professor for assembly and packaging with the Department of Microsystems Engineering, University of Freiburg, Freiburg, Germany. The focus of his work lies in electronic packaging technologies and their effect on reliability and functionality.

Prof. Wilde is the Chairman of the "Electronics Packaging" board of the VDE/VDI Society GMM.

Peter Alpern received the Ph.D. degree in physics from the Technical University of Munich, Munich, Germany.

He was with Siemens AG in 1984. In 1987, he transferred to Siemens Semiconductor Division. He is currently with Infineon Technologies AG, Neubiberg, Germany, where he is engaged in quality/reliability development for the overall system chip and package.

Dr. Alpern is member of the German Physical Society.

Matthias Stecher received the M.S. degree from the Virginia Polytechnic Institute and State University, Blacksburg, and the Ph.D. degree from Rheinisch-Westfälische Technische Hochschule, Aachen, Germany, in 1994, both in electrical and electronic engineering.

Between 1998 and 1994, he was involved in the development of device and circuit simulation tools. Since 1994, he has been with Infineon Technologies AG, Neubiberg, Germany, where he has been the Project Manager for several Smart Power Technologies. Since 2003, he has been involved in the thermomechanical optimization of chip-package systems and is currently a Technical Advisor in the fields of power technology and package development. He is the holder of more than 20 patents.