Finite-Element Analysis of Coupled Electro-Thermal Problems With Strong Scale Separation

Sebastian Eiser, Mirko Bernardoni, Michael Nelhiebel, and Manfred Kaltenbacher, *Member, IEEE*

***Abstract*—We present a finite-element analysis of the coupled electro-thermal behavior of a mm^2-sized power DMOS device, with a focus on resolving µm-sized structures. For a reliable power technology, the predictions of current and temperature distributions under self-heating are vital for reaching long device lifetime. Two aspects of our approach are unique: First, we have integrated the active region of the DMOS into the finite-element program code as a nonlinear voltage and temperature dependent *tripole*, incorporating local biasing and temperature conditions. Second, to model small features like defects or process-induced imperfections in sufficient detail, we use the *Nitsche*-type nonmatching grid technique to include the substantially finer meshed microscale geometry. This allows us to study very local changes in temperature and biasing conditions, which are normally not accessible. The methodology is applied to a state-of-the-art power technology device with integrated current sensor to validate the methodology. Good agreement is found between results and experiments.**

***Index Terms*—DMOS, electro-thermal effects, finite-element methods, heat transfer, multiscale modeling, power semiconductor devices.**

I. Introduction

ADVANCES in chip manufacturing technologies and shrinkage of active device area are leading to increasing power density in semiconductor devices. As a result, more effort is spent on thermal management and design optimization to achieve high device reliability in the field. Today's automotive power devices are capable of withstanding $1300\,\mathrm{W/mm^2}$ pulses [1] in repetitive load conditions. However, care must be taken both to prevent junction overtemperature [2] and excessive, thermally induced strain caused by the mismatch in the coefficient of thermal expansion and/or strong temperature gradients [3]–[5]. In order to achieve a satisfactory lifetime performance [6], [7], effective analysis tools are required to support engineering decisions. Frequently, the driving failure mechanism of a known point-of-failure is sought after. Unfortunately, many problems are difficult to analyze efficiently due to the inherent scale separation or high-aspect layers present in semiconductor devices. For example, µm-sized imperfections in metalization, defects in trenches (e.g., field oxide) resulting from assembly, process variation, or front-end production. These may change the *local* electrical properties (e.g., *threshold voltage*), or cause small-scale inhomogeneities in the temperature field. Furthermore, due to the piezo-electric effect in *GaN*-based devices, local impurities may significantly impact device performance. These problems are most commonly ignored due to the inaccessibility of an analysis resolving both the small features *and* the surroundings.

In traditional analysis schemes (e.g., *RC* networks [8], [9], lumped element models [10], [11], finite differences [12] or finite volumes [13]) small complex geometries capturing electro-thermal effects are very cumbersome to integrate. The strength of the finite element (FE) method is the flexible modeling of arbitrarily shaped topologies. Yet, conventional FE analysis also fails to efficiently solve multiscale problems due to the penalty incurred by the high amount of degrees of freedom arising in traditional, conformal FE meshes [14], [15]. This causes delays in efficient pre and postprocessing, as well as in solution time. For example, accurately resolving the thin heat generation film (µm-scale) in a power DMOS (few mm^2) is limited by the maximum aspect ratio of FE elements required for an accurate solution. Complex geometries are therefore normally simplified based on the very specific problem.

In Fig. 1(a), the mesh of a semiconductor device is shown with one region of interest highlighted. Note how the requirement of fine meshing in one region requires coarse meshing in adjacent regions where homogenized material parameters are desirable and a lower spatial resolution could be possible (both in lateral and vertical direction). Attempts to discretize thin films, with a fast transition zone, frequently result in unsuitable multiscale meshes [see Fig. 1(b)].

To analyze the problem efficiently, a very fast transition from small to coarse element size is necessary. Nonmatching interfaces greatly simplify this procedure of including small-scale fully resolved geometries into the modeled device [see, e.g., Fig. 1(c)]. In addition to the greatly simplified meshing effort, the total mesh size is also reduced, which facilitates pre and postprocessing, and reduces simulation time.

We have already investigated the *Mortar* FE method [16] for coupling subdomains with independent meshes and occasionally found unphysical oscillations on the interface (possibly caused by an inappropriate intersection mesh). In this study,

Manuscript received August 5, 2015; revised December 20, 2015; accepted January 27, 2016. Date of publication February 11, 2016; date of current version September 16, 2016. This work was supported in part by the Austrian Research Promotion Agency (FFG, Project 846579) and in part by the Carinthian Economic Promotion Fund (KWF, Contract KWF-1521-26876-38867). Recommended for publication by Associate Editor H. Wang.

S. Eiser is with the KAI Competence Center of Automobile and Industrial Electronics, Villach 9524, Austria, and also with the Faculty of Mechanical and Industrial Engineering/Institute of Mechanics and Mechatronics, Vienna University of Technology, Vienna 1000, Austria (e-mail: sebastian.eiser@student.tuwien.ac.at).

M. Bernardoni is with the KAI Competence Center of Automobile and Industrial Electronics, Villach 9524, Austria (e-mail: mirko.bernardoni@k-ai.at).

M. Nelhiebel is with the KAI Competence Center of Automobile and Industrial Electronics, Villach 9524, Austria, and also with the Infineon Technologies Austria AG, Villach 9500, Austria (e-mail: michael.nelhiebel@infineon.com).

M. Kaltenbacher is with the Vienna University of Technology, Vienna 1000, Austria (e-mail: manfred.kaltenbacher@tuwien.ac.at).

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/TPEL.2016.2527690

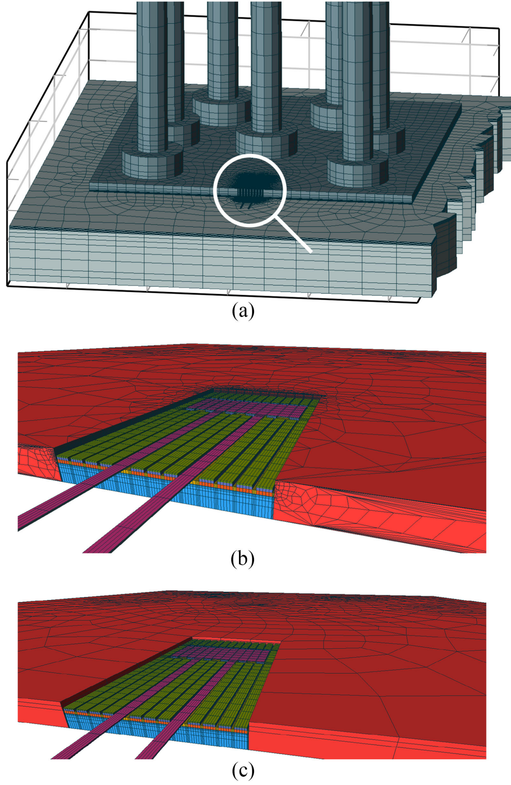

Fig. 1. Comparison of meshing strategies: (a) a typical model and mesh with silicon substrate and bond wires (both partially clipped). The mesh is selectively refined to resolve a region of interest. Due to the transition from fine mesh to coarse mesh, many degrees of freedom are required with a classical meshing strategy. (b) a failed attempt to mesh with a fast transition from fine to coarse element size. (c) a nonmatching discretization with independent subdomain meshes.

we are using *Nitsche*-type mortaring to include a finely meshed structure into a coarse chip model in a strongly coupled electro-thermal FE simulation. The Nitsche method penalizes differences in temperature or electric potential on the nonmatching interface, and ensures the correct heat flux and electric current, respectively, across the interface, an important property for bidirectionally coupled problems.

This paper is organized as follows: Background and theory of the electro-thermal FE formulation are outlined in Section II and details on the implementation of the DMOS and the nonmatching method are given. In Section III, the framework is applied to the modeling of a real application: a power DMOS with its current sensor resolved in detail, serving as an experimentally accessible test bed. Results of these investigations as well as experimental comparisons are given in Section IV. Lastly, the results are summarized and a conclusion is provided.

II. Background and Coupled Formulation

A. Physical Equations

The physical equations of the thermal and electrical behavior are the *heat equation*

$$\rho c \frac{\partial T}{\partial t} - \nabla \cdot (\lambda \nabla T) = p_{\mathrm{V}} \tag{1}$$

the *charge continuity equation* in matter

$$\nabla \cdot \vec{j} = \nabla \cdot (\gamma \nabla \phi) = 0 \tag{2}$$

and their mutual interaction. In (1), ρ is the density, c the specific heat capacity, λ the thermal conductivity, and T the temperature. Furthermore, p_{V} represents the heat generation rate per unit volume. In (2), $\vec{j}$ is the current density, γ is the electric conductivity, and ϕ is the electric potential. We assume that the displacement current density $\vec{j}_D$ caused by polarization of bound charges in the dielectric can be neglected in our application. In addition, the conductivities λ and γ may depend on the flux direction ($\vec{q} = \tilde{\lambda}\nabla T$ or $\vec{j} = \tilde{\gamma}\nabla\phi$, respectively). In this case, a tilde (e.g., $\tilde{\lambda}$) distinguishes between anisotropic tensorial material parameters and scalar material parameters. The coupling of the electric partial differential equation (PDE) (2) into the heat equation (1) is given by means of joule heating

$$\begin{aligned} p_{\mathrm{V}} &= \vec{E} \cdot \vec{j} \\ &= (\nabla \phi) \cdot (\gamma \nabla \phi) \end{aligned} \tag{3}$$

where $\vec{E}$ is the electric field. In turn, the heat equation causes changes in the temperature distribution $T(\vec{x})$ which affects the electric conductivity $\gamma(T)$. *Dirichlet* boundary conditions (BCs) for the heat and electrical conduction are used to impose temperature T and electric potential ϕ on (parts of) the domain boundary Γ_{e}. *Von Neumann* BCs may be used to impose heat flux and electrical current, respectively.

B. Electro-Thermal FE Formulation

Upon discretizing the temporal and the spatial domain, the resulting algebraic system of the weak form of (1) and (2) [16] can be expressed in matrix notation

$$\mathbf{M}(\underline{T}_{n+1})\,\dot{\underline{T}}_{n+1} + \mathbf{K}^{T}(\underline{T}_{n+1})\,\underline{T}_{n+1} = \underline{f}_{T,n+1} \tag{4}$$

$$\mathbf{K}^{\phi}(\underline{T}_{n+1})\,\underline{\phi}_{n+1} = \underline{f}_{\phi,n+1} \tag{5}$$

where $\mathbf{M}$ is the mass matrix, $\underline{f}_{T,n+1}$ is the right-hand side load vector resulting from the power dissipation p_{V} (3) and the index n denotes the time-step counter. $\mathbf{K}^{T}$ and $\mathbf{K}^{\phi}$ are the stiffness matrices for the thermal and electrical problem, and $\underline{f}_{\phi,n+1}$ is the right-hand side of the electric problem arising from *von Neumann* BCs. $\dot{\underline{T}}_{n+1}$, $\underline{T}_{n+1}$, and $\underline{\phi}_{n+1}$ are the algebraic vectors of time derivative of temperature, temperature, and electric potential defined on the nodes of the mesh, respectively. Time integration is performed according to the backward Euler scheme [17]. For further details see [17]–[19]. The solution strategy for linear material parameters is already detailed in [16].

Both (1) and (2) are nonlinear and can be solved using an iterative solution scheme. Fig. 2 shows the solution algorithm in which the iteration counters i, k, and l are introduced. k and l are incremented at each iteration of solving the nonlinear heat and electric equation by means of a *Fixed-point* method. These inner iterations use an incremental stopping criterion ($u = \{T, \phi\}$, respectively) for the new time step $n + 1$

$$\frac{\left\|\underline{u}^{m+1} - \underline{u}^{m}\right\|_2}{\left\|\underline{u}^{m+1}\right\|_2} < 10^{-3} \tag{6}$$

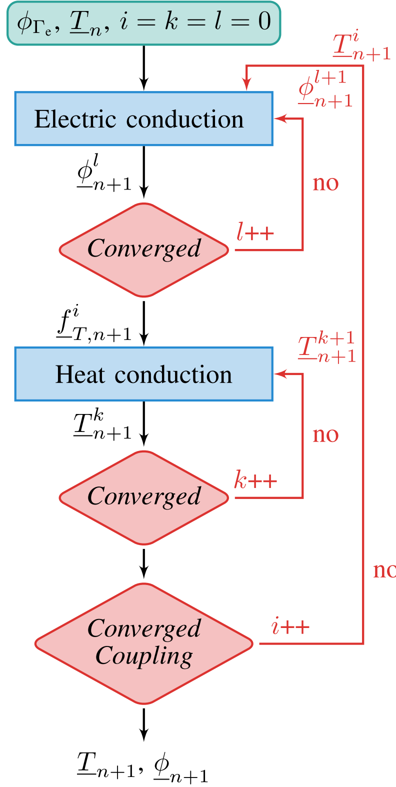

Fig. 2. Iterative convergence strategy for the coupled electro-thermal problem for a single time step n. The load vector $\underline{f}^i_{n+1}$ of the nonlinear heat equation is calculated from the power density of the electric conduction. The outer iteration (i) ensures the convergence of the coupled problem.

where $\|\cdot\|_2$ is the L_2-norm and $m = \{k, l\}$ the iteration counter. Furthermore, a residual-based stopping criterion has to be fulfilled as well

$$\frac{\left\|\underline{f}^i_u - \mathbf{M}(\underline{u}^{m+1})\underline{\dot{u}}^{m+1} - \mathbf{K}^u(\underline{u}^{m+1})\underline{u}^{m+1}\right\|_2}{\left\|\underline{f}^i_u\right\|_2} < 10^{-3}. \quad (7)$$

Note that, for $u = \phi$, we have $\underline{\dot{u}} = 0$. The outer iteration counter i is incremented at each coupled iteration, corresponding to the overall convergence between the electrical and thermal equation. For each new time step $n+1$, the following iterative solution scheme is performed:

1) Set $\underline{\phi}^{l=0}_{n+1} = \underline{\phi}^i_{n+1}$ and iteratively solve the *inhomogeneous* and *nonlinear* electric conduction problem (5) with a given temperature $\underline{T}^i_{n+1}$ Note that, initially $\underline{T}^{i=0}_{n+1} = \underline{T}_n$. For each iteration l, solve

$$\mathbf{K}^\phi(\underline{T}^i_{n+1})\,\underline{\phi}^{l+1}_{n+1} = \underline{f}_{\phi,n+1}. \quad (8)$$

Upon fulfilling the stopping criteria (6), (7), the distribution of the electric potential $\underline{\phi}^{i+1}_{n+1}$ is obtained.

2) Compute the temperature-dependent joule heating term $\underline{f}^i_{T,n+1}(\underline{\phi}^{i+1}_{n+1}, \underline{T}^i_{n+1})$ using (3).

3) Set $\underline{T}^{k=0}_{n+1} = \underline{T}^i_{n+1}$ and iteratively solve the *nonlinear* heat equation ((4), inner k loop)

$$\mathbf{M}(\underline{T}^k_{n+1})\underline{\dot{T}}^{k+1}_{n+1} + \mathbf{K}^T(\underline{T}^k_{n+1})\underline{T}^{k+1}_{n+1} = \underline{f}^i_{T,n+1}. \quad (9)$$

After reaching convergence (6) and (7), $\underline{T}^{i+1}_{n+1}$ is obtained.

4) Check for the convergence of the coupled systems, which is based on the change of the joule heating term between two subsequent outer iterations

$$\left\|\underline{f}^{i+1}_T - \underline{f}^i_T\right\|_2 < 10^{-3}\cdot\left\|\underline{f}^{i+1}_T\right\|_2. \quad (10)$$

If convergence is achieved, we set

$$\underline{T}_{n+1} := \underline{T}^{i+1}_{n+1} \text{ and } \underline{\phi}_{n+1} := \underline{\phi}^{i+1}_{n+1} \quad (11)$$

otherwise we repeat the steps from 1 to 4.

Direct coupling of (1) and (2) via (3) can be achieved by using a Newton scheme. We chose to solve the coupled system sequentially as shown in Fig. 2. This approach is advantageous in our case since only a small number of iterations is required until convergence is reached.

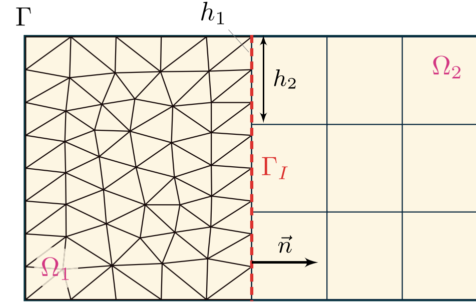

Fig. 3. Computational domain Ω has been decomposed into two independently meshed subdomains Ω_1 and Ω_2. On the subdomain interface Γ_I, the mismatch in discretization is handled by additional surface integrals.

C. Nonmatching Grids

The regular FE method has strict requirements on the discretization of the computational domain Ω, i.e., the input mesh. In two dimensions, adjacent elements must share only one edge and in three dimensions only one surface. Also, the elements must retain a good aspect ratio to allow accurate calculation of the Jacobian determinant. The FE program usually checks their shape by means of quality criteria.

In our previous work [16], the *Mortar* FE method [20], [21] has been utilized as a nonmatching domain decomposition technique, removing the above limitation on the number of adjacent elements. The main disadvantage of the *Mortar* method is the arising saddle point structure of the resulting linear system of equations, requiring sophisticated numerical solver packages [22]. A possibility to circumvent this drawback would be the use of the dual Lagrange multiplier method [21], [23], which is however quite demanding concerning the computer implementation. Therefore, we have investigated *Nitsche*-type mortaring [24], a nonmatching domain decomposition FE method. In contrast to the *Mortar* method, neither the resulting linear system is changed, nor new unknowns are introduced (e.g., *Lagrange* multipliers). *Nitsche's* approach works by adding additional, symmetrizing surface integrals on Γ_I, and a penalty term to the weak formulation. Fig. 3 displays two independently meshed subdomains (Ω_1, Ω_2) joined at a nonmatching interface Γ_I. Here, we demonstrate this method for the time-independent

Poisson problem with scalar unknown u, in the computational domain $\Omega = \Omega_1 \cup \Omega_2$

$$\begin{aligned} -\nabla k(\vec{x})\nabla u &= f \qquad \text{in } \Omega \\ u &= 0 \qquad \text{on } \Gamma = \partial\Omega. \end{aligned} \tag{12}$$

In (12), k is the possibly nonlinear material parameter. For simplicity, k_1 and k_2 are assumed to be constant in each subdomain Ω_1 and Ω_2, but this is by no means a restriction. The weak form for subdomain i reads

$$\int_{\Omega_i} k_i \nabla w_i \cdot \nabla u_i \, \mathrm{d}\Omega - \int_{\Gamma_I} k_i w_i \frac{\partial u_i}{\partial \vec{n}_i} \, \mathrm{d}\Gamma = \int_{\Omega_i} w_i f_i \, \mathrm{d}\Omega \tag{13}$$

where $\vec{n} = \vec{n}_1 = -\vec{n}_2$ is the unit normal vector pointing from Ω_1 to Ω_2, as shown in Fig. 3 and w is a suitable basis function. After summing over $i = \{1, 2\}$, performing symmetrization and adding a penalty term which ensures that the value of the unknown u is equal on the interface, we arrive at a coupled formulation (see [17] and references therein)

$$\begin{aligned} \sum_i \int_{\Omega_i} k_i \nabla w_i \cdot \nabla u_i \, \mathrm{d}\Omega \; - \int_{\Gamma_I} k_i [w] \frac{\partial u_1}{\partial \vec{n}} - k_i \frac{\partial w_1}{\partial \vec{n}} [u] \, \mathrm{d}\Gamma \\ + \beta \overline{k} \int_{\Gamma_I} [w][u] \, \mathrm{d}\Gamma = \sum_i \int_{\Omega_i} w_i f_i \, \mathrm{d}\Omega. \end{aligned} \tag{14}$$

In (14), the jump operator $[u] := u_1 - u_2$ has been introduced. Furthermore, $\overline{k} := (k_1 + k_2)/2$ and β is the mesh-dependent penalty factor. To remove the mesh dependency, we choose $\beta_0 = 20$ and used the following relation for a robust method:

$$\beta = \beta_0 \frac{\max(p_1, p_2)^2}{\min(h_1, h_2)} \tag{15}$$

where h_i and p_i are element size and order of the basis function on either side of the interface (see Fig. 3).

After discretization, the resulting system matrix is

$$\left[\overbrace{\begin{pmatrix} K_{11} & 0 \\ 0 & K_{22} \end{pmatrix}}^{\text{indep. subdomains}} + \overbrace{\begin{pmatrix} K_{\Gamma_1\Gamma_1} & K_{\Gamma_1\Gamma_2} \\ K_{\Gamma_2\Gamma_1} & K_{\Gamma_2\Gamma_2} \end{pmatrix}}^{\text{coupling \& penalty}} \right] \begin{pmatrix} \underline{u}_1 \\ \underline{u}_2 \end{pmatrix} = \begin{pmatrix} \underline{f}_1 \\ \underline{f}_2 \end{pmatrix} \tag{16}$$

where $K_{\Gamma_i\Gamma_j}$ arises from the surface integrals in (14).

In addition to the thermal and electrical problems we are discussing in this study, *Nitsche*-type mortaring has been also applied to fluid-structure interaction [24] and contact mechanics [25].

D. DMOS Electro-Thermal Modeling

The temperature dependent DMOS transconductance is

$$I_D = f(V_{\mathrm{GS}}, V_{\mathrm{DS}}, T) \tag{17}$$

where I_D is the drain current, V_{GS} is the gate–source voltage, and V_{DS} is the drain–source voltage. For a DMOS operating in the saturation region [26], i.e., $V_{\mathrm{DS}} > V_{\mathrm{GS}} - V_{\mathrm{th}} > 0$

$$I_D = \mu_n(T) C_{\mathrm{ox}} \frac{W}{2L} \left(V_{\mathrm{GS}} - V_{\mathrm{th}}(T) \right)^2. \tag{18}$$

In (18), channel width W, channel length L, and device specific technology parameters such as oxide specific capacity C_{ox}, temperature-dependent electron mobility $\mu_n(T)$, and threshold voltage $V_{\mathrm{th}}(T)$ are given. The temperature induced change in the drain current I_D in (18) is given by the dependency of $\mu_n(T)$ and $V_{\mathrm{th}}(T)$. Typically, a *stable* and *unstable* regime exist, separated by the *temperature compensation point* (TCP) $V_{\mathrm{GS}} = V_{\mathrm{TCP}} > V_{\mathrm{th}}$ [27]. The unstable regime is characterized by the positive feedback of heat on the drain current ($\partial I_D / \partial T > 0$). Note that temporary operation in the unstable region can never be avoided when switching OFF the device. In the case of inductive loads, the demagnetization times may result in substantially long operation below the TCP.

Because of the large active area of a power DMOS, the spatial inhomogeneity of temperature may be significant. As a result, the operation of a power DMOS below the TCP will cause current crowding in temperature hot spots. In this case, the active area can be laterally decomposed into a sufficient amount of cells for which the model in (17) is still approximately valid by using $I_D^i = f(V_{\mathrm{GS}}^i, V_{\mathrm{DS}}^i, T_i)$ for each cell i.

E. Implementation of the DMOS Model in the FE Formulation

In the FE formulation, the $I_D\,(V_{\mathrm{GS}}, V_{\mathrm{DS}}, T)$ characteristic of the DMOS is modeled by the electrical conductivity γ of the transistor region

$$\gamma(\phi, T) = \gamma(V_{\mathrm{GS}}, V_{\mathrm{DS}}, T) \tag{19}$$

$$V_{\mathrm{GS}} = \phi_G - \phi_S \tag{20}$$

$$V_{\mathrm{DS}} = \phi_D - \phi_S. \tag{21}$$

This is a *nontypical* nonlinearity, because γ depends on the *difference* of the electrical potential ϕ. A material law for a temperature dependent *tripole* is required, which is not typically implemented in FE programs (usually, $\gamma(\phi, T)$ is available).

In the work of de Filippis *et al.* [28], this problem has been addressed by performing a preliminary, nonlinear electrical iteration with nonuniform temperature distribution $T(t_{n-1}, \vec{x})$, to set the channel conductivity $\gamma(V_{\mathrm{GS}}, V_{\mathrm{DS}})$. Subsequently, the coupled electro-thermal system is solved with this value of γ. See also [29] for details on ensuring the correct power dissipation and convergence criteria.

In our approach, no preliminary step is required due to the full access to the FE research code of [22], in which the custom, temperature-dependent tripole has been implemented as a new material law. Thus, temperature and electric potential distribution are more finely resolved (down to *Gauss* integration points) and efficient coupling of the PDEs is possible. The directional dependency of the current flow in a vertical DMOS is accounted for by means of an anisotropic tensorial conductivity $\tilde{\gamma}(V_{\mathrm{GS}}, V_{\mathrm{DS}}, T)$.

The main advantage of our approach and [28] is the incorporation of the spatial inhomogeneity of the electric potential (debiasing) and the temperature field (see also [30]). The DMOS model used in this study was generated from TCAD simulations and provided as an $I_D = f(V_{\mathrm{GS}}, V_{\mathrm{DS}}, T)$ lookup table.

Fig. 4. Schematic cross-section of the modeled structure (not to scale). Die attach and leadframe must be included in the model given the time duration of the power pulse. Due to its low thermal conductivity and capacity and the short time scales (< 2 ms), packaging impact is negligible.

Fig. 5. Equivalent circuit and measurement schematics of the device. The sensor and the main DMOS form an unbalanced current mirror with high area ratio $k_{\text{geo}} \gg 1$. Due to the size of the larger DMOS', temperature and source voltage may exhibit significant spatial inhomogeneity (indicated by angle brackets $\langle\cdot\rangle$). R_S and R_{int} are internal parasitic resistances. The *Zener* diode V_Z and R_Z are required for *clamped* measurements.

III. Model Development

For validation purposes, we chose to test the new methodology by means of a trench power DMOS [1] with an integrated current sensor. The availability of the sensor allows us to:

1) Perform a *proof-of-concept* of an experimentally accessible bidirectionally coupled electro-thermal problem.
2) Validate the electrical and thermal performance of the nonmatching interface.

A. Test Device Description and Modeling Approach

The device is a typical low-voltage power device consisting of several thousand parallel elementary cells encapsulated in a plastic package with exposed lead frame (drain) in Fig. 4. The exposed lead frame of the vertical device and the short operating times result in negligible impact of modeling the plastic package or not. The sensor current is accessible through an extra pin on the package. Fig. 5 shows the equivalent circuit of the device under test. In our case, the current mirror is intentionally unbalanced via a shunt resistor R_{shunt}. Normally, these parallel transistors are operated in *balanced* mode, i.e., an *operational amplifier* ensures that $\phi_S^{\text{DMOS}} \approx \phi_S^{\text{sense}}$. We choose a bare device to control the degree of debiasing with the shunt resistor R_{shunt}, which also allows us to measure the sensor current I_{sense}.

Fig. 6. Example of a nonmatching interface of a hypothetical sensor geometry. (a) top view on the active area (A-A view) (b) cross section with material stack on the substrate top side at the B-B slice. The z-axis has been rescaled to facilitate visualization.

In Fig. 6(a), a top view of the nonmatching interface separating the coarse, homogenized model from the finely resolved, hypothetical sensor geometry is shown, and in Fig. 6(b), the cross-section is given with mesh omitted for clarity. The true sensor looks similar and is functionally equivalent. In the coarse model, anisotropic homogenized material parameters are used for the inter layer dielectric (ILD) and the active DMOS layer. The material properties are identical to the ones given in an earlier publication [16]. The DMOS layer is approximated by a single layer of elements with thermal properties like the epitaxial silicon layer. However, due to its small z-extension with only one element and the fact that the heat generation takes place in this region, the thermal properties are not playing a significant role. The electrical conductivity is only nonzero for the z-direction to account for vertical current flow. This allows us to model the current distribution in the DMOS plane more accurately in the case of spatially inhomogeneous drain or source potentials. The value of the z-conductivity $\gamma_z(V_{\text{GS}}, V_{\text{DS}}, T)$ is set in every iteration of the electric PDE according to the scheme outlined in Section II-E.

The resistors R_{int} and R_{shunt} (see Fig. 5) are modeled as volumes in the FE simulation [see extrusions in Fig. 1(c)].

B. FE Simulation and BCs

The main uncertainties lie in the approximation of the model, its BCs and the selected material parameters. In particular, the thermal and electrical nonlinearities in the silicon epitaxial and silicon substrate layers are challenging to address. We use the previously validated material model for the silicon substrate (nonlinear) and homogenized intermetal-dielectric (anisotropic) as in our earlier work [16]. The DMOS model from TCAD has previously been validated over a broad range of operating conditions by means of a finite volume simulation and experiments [13].

1) Thermal BCs: Accurate modeling of the thermal BCs is important in order to obtain the correct power distribution in the active area of the device, given the strong electro-thermal interaction. Due to the presence of the time derivative in the PDE (1), element size and temporal step size are connected. When using small time steps, elements in the thermal mesh must be sufficiently small to prevent unphysical results. Furthermore, the thermal mass of the layers below the substrate (e.g., leadframe, PCB, etc.) must be included depending on the length of the analysis. The longer the time duration of the power pulse, the more layers need to be included in the model. We deal with pulses in the range of milliseconds and heuristically found that including die-attach and lead-frame volumes correctly captures the heat propagation in the structure. We use approximately 25×10^3 and 5×10^3 linear volume elements for the macro and micro model, respectively. Only one corner at the bottom of the leadframe is kept at fixed ambient temperature $T_{\text{amb}} = 25\,^\circ$C, for numerical reasons (see Fig. 4). The entire model is set to ambient temperature as initial condition.

2) Electrical BCs: In the model (cf., Fig. 4), the measured drain–source voltage V_{DS} is applied between substrate and bond wires (ground). The voltage drop across the highly doped substrate is neglected. On the source side, however, high currents will lead to *debiasing*, causing a spatial inhomogeneity in V_{GS} and V_{DS} and, hence, a laterally inhomogeneous operating point for each integration point [30]. This is taken into account in our method. The gate voltage V_{GS} is read from a table of measurement data. The additional layers below the substrate, required for correctly modeling the heat propagation, have no influence on the electrical behavior of the device. Therefore, the electrical problem does not have to be solved in these regions, considerably improving solution speed.

IV. Results and Discussions

The primary results of the coupled simulations are the electric potential $\phi(\vec{x}, t)$ and temperature field $T(\vec{x}, t)$. From these, subsequent results such as electrical current density, power dissipation or thermal fluxes can be obtained.

A. Electro-Thermal Model Validation

At first, the DMOS model in the FE simulation (Section II-E) and the model assumptions (Section III) have to be validated. To this end, experimental data are required in order to compare predictions of the simulation to transient electrical measurements. Furthermore, the destruction of a device in thermal run-away occurs at a well-known temperature T_{runaway}. Therefore, for a given pulse, agreement in destruction time and melt-up location are ideal validation instruments. The measurement setup of the active *Zener*-clamped device is shown in Fig. 5. Experiments are performed at room temperature (25 °C). Current pulses are supplied to the device using an external programmable current source, set to operate the devices below the TCP. Under this load, device failure is expected due to thermal run away in the hottest region of the active area [27].

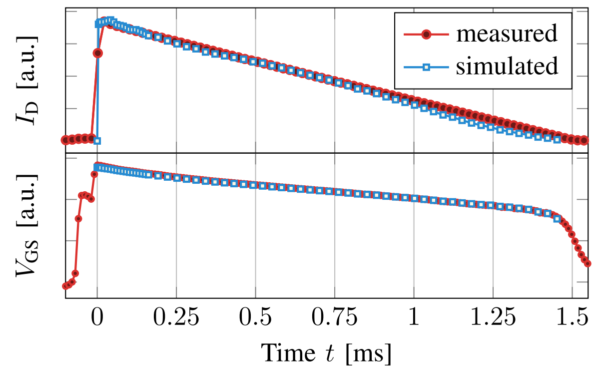

Fig. 7. Comparison of measured versus simulated results. The simulation is forced to match the measured current I_d^{exp} within a defined tolerance, to ensure the correct power dissipation. The simulation is then validated by comparison the output of the gate–source voltages V_{GS}.

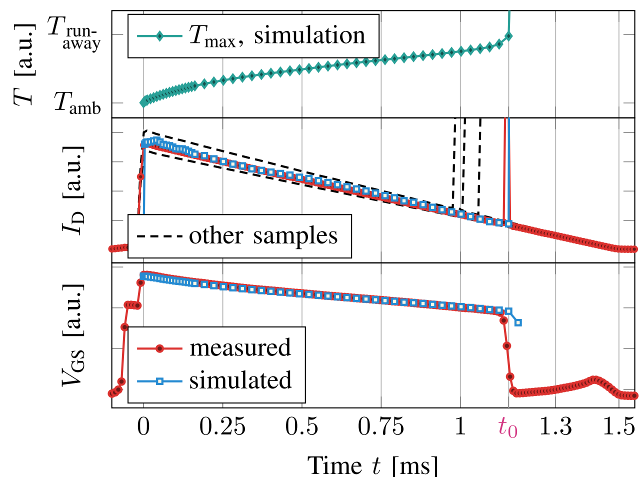

Fig. 8. Comparison of measured versus simulated results for a destruction pulse. Device variation in the destruction current is indicated by the dashed lines (measurement of three other samples). Operation below the TCP causes current crowding and melt-up in the hot spot (cf., Fig. 9).

A comparison of measured and simulated results is shown in Fig. 7. The *simultaneous* agreement of V_{GS} and I_D is a very good indication that both the DMOS model $\gamma(\phi, T)$ as well as the temperature field in the device are simulated with sufficient accuracy. The high temperature rise in the device for this pulse of $\Delta T_{\text{max}} \approx 400$ °C and resulting strong electro-thermal interaction indicate the validity of material models and model approximations. No fitting parameters are required to obtain this result. A set of destruction current pulses is shown in Fig. 8. Due to manufacturing process tolerances (e.g., layer thicknesses),

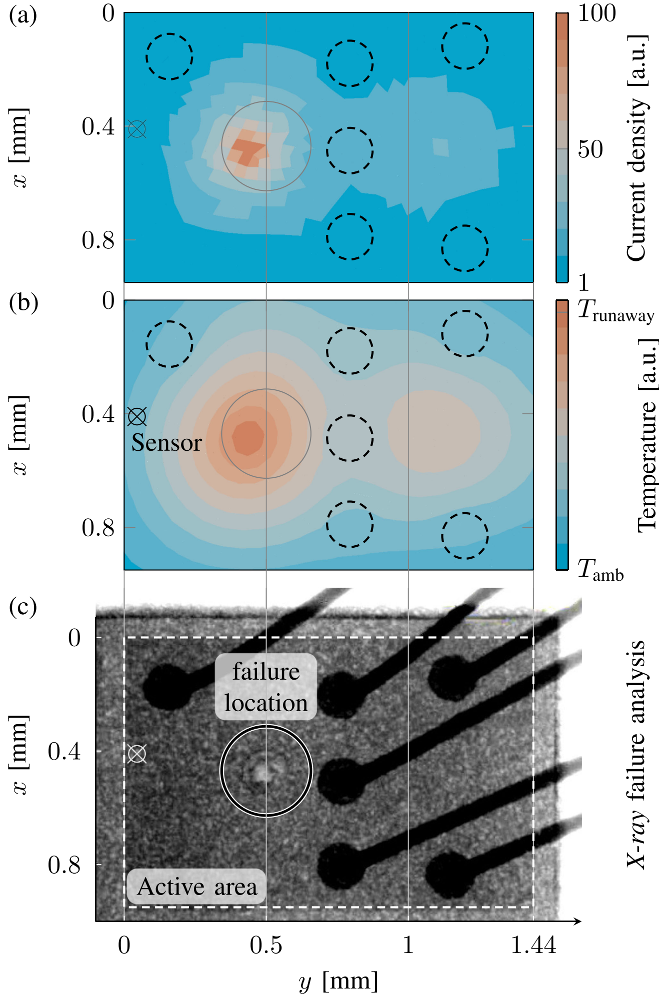

Fig. 9. Operation below the TCP at $t_0 = 1.15$ ms for the destruction pulse: The maxima of the electric current density (a) and temperature distribution (b) of the active region coincide, shortly before run away. The nailhead positions are indicated by dashed circles. In (c), the location of the *melt-up* is visible in an contrast enhanced *X-ray* tomogramm by a bright crater (circled) close to the predicted failure site (b).

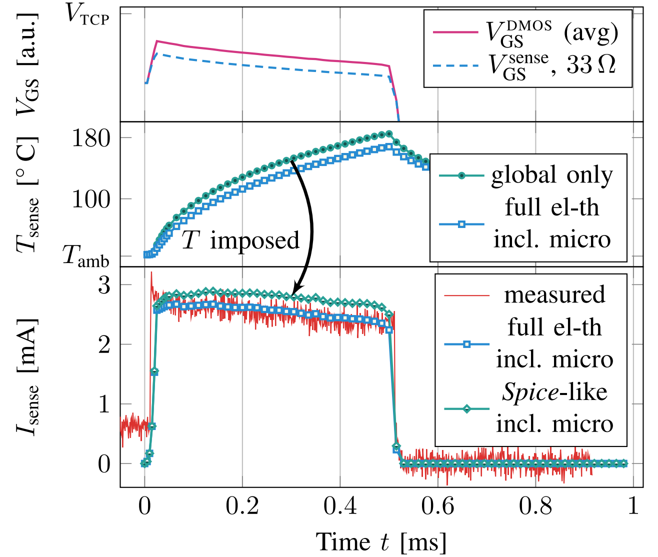

Fig. 10. *Top:* The shunt resistor ($R_{\text{shunt}} = 33\,\Omega$) reduces the local gate–source voltage at the sensor DMOS cells. *Middle:* The temperature in the sensor center is overestimated in the global model due to an insufficiently resolved sensor geometry. *Bottom:* Comparison of a *Spice*-like simulation (unidirectional thermo-electric coupling in the sensor DMOS) and the fully coupled model with the sensor resolved. The imposed, higher temperature causes an 11% higher sensor current.

destruction energies follow a probability distribution, visualized in Fig. 8 with four destruction events.

The current-crowding phenomenon and hot-spot formation can be observed by looking at the temporal evolution of the current density and the temperature field in the active region of the DMOS. Fig. 9 shows the simulated spatial inhomogeneity of $j_D(\vec{x})$ and $T(\vec{x})$ shortly before run-away occurs. As seen in Fig. 9(a), the current density is highest in the hottest region. The temperature distribution in Fig. 9(b) shows the cooling effect of the nailheads and bond wires, causing the hot spot to develop in the left section of the active region. In Fig. 9(c), an *X-ray* tomograph of the device shows the melt-up location, which confirms the predicted failure location of the simulation. We, therefore, have gained confidence in the adequacy of the electro-thermal coupling scheme and the transistor model applied in this study.

B. Sensor Current and Current Ratio

With the new methodology we have the opportunity to analyze the sensor operating at different biasing conditions and check the performance of the coupled, nonlinear simulation including the nonmatching grid. In this setup, the sensor current I_{sense} enters the micro model through the nonmatching interface on the drain side [cf., Fig. 6(b)]. Due to the debiasing of the sensor by means of R_{shunt}, we expect I_{sense} to decrease with increasing R_{shunt}. However, quantifying the decrease is difficult because of the electro-thermal interaction. A lower sensor current reduces the local heat generation which in turn shifts the sensor's operating point, again affecting I_{sense}. The sensor is thermally fully connected to the device by means of the nonmatching interfaces surrounding it, permitting heat exchange. The topology of the sensor further increases the difficulty in forecasting its local temperature. On one hand, its thermal insulation from the large heat capacity on top [cf., Fig. 6(b)] could cause elevated temperatures for high power dissipation and short time scales (see [16]). On the other hand, with its positioning close to the boundary of the DMOS, the cooling effect may dominate. Since I_{sense} strongly depends on the temperature, it can be used to validate its correct thermal evolution. Initially, the simulated current was lower than the measured value, independent on R_{shunt}, which can be mitigated by increasing the conductivity of the sensor cells by 30%. Tests have been performed to exclude the nonmatching interface as the source of this mismatch. Neither jumps in electric potential or temperature on the interface nor discontinuities in the electrical current or heat flux were observed. Since the value is independent on R_{shunt} , process-induced differences in the electrical characteristics of the sensor region with respect to the DMOS or parasitic resistances may be responsible for this discrepancy. The increase in the conductivity is the only fitting parameter in our model.

In Fig. 10, the debiasing of the sensor DMOS is visualized during a rectangular current pulse with $R_{\text{shunt}} = 33\,\Omega$. The negative offset in the gate–source voltage causes proportionally less

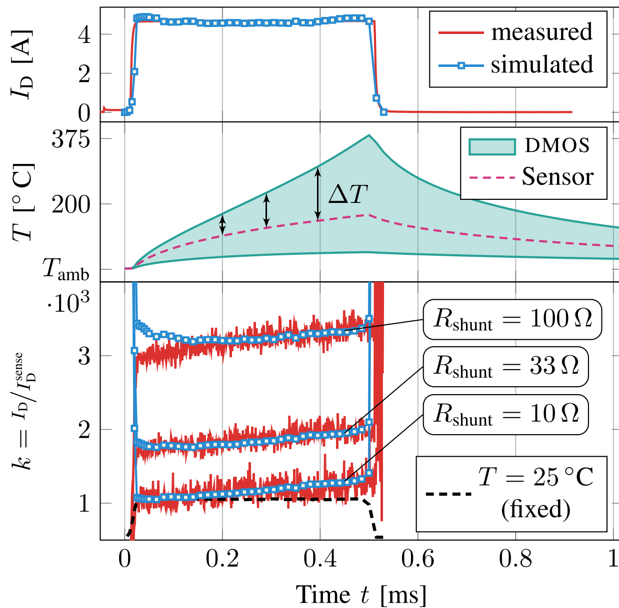

Fig. 11. Comparison of measured versus simulated results for the sensor current ratio for three different values of the shunt resistor R_{shunt}. *Top*: Total current; *Middle*: Max/Min temperature of the DMOS and the sensor; *Bottom*: Sensor current ratio in real operation and simulated *isothermal* case (dashed line).

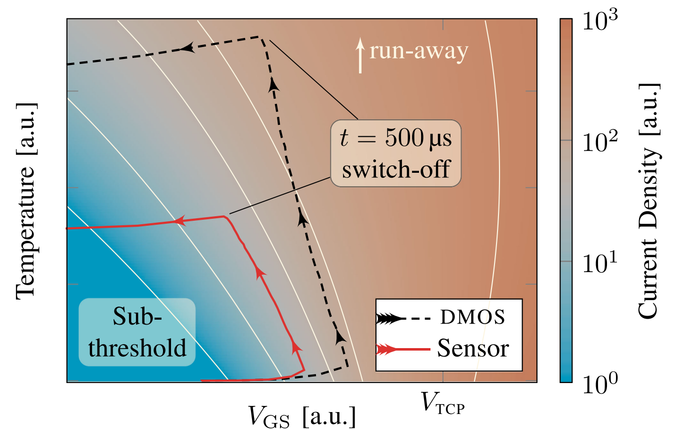

Fig. 12. The temperature maxima of the main DMOS and the debiased sensor ($R_{shunt} = 33\,\Omega$) are superimposed as paths on the transistor transconductance for the $500\,\mu s$ pulse shown in Figs. 10 and 11. The divergence of the two paths indicates the shift of the main DMOS towards run-away and higher conductance.

current in the sensor. The decrease in both gate–source voltages is caused by operating the device below the TCP: As the device heats up by more than 300 K, a lower V_{GS} is required to sustain the constant current [27]. For comparison, we also include the case of the *global* model, i.e., the model from the previous section, which neither features the nonconforming mesh nor the sensor debiasing (filled circles in the middle plot of Fig. 10). The overestimation of the temperature originates both from efficient cooling of the sensor DMOS by the metal one wire present in the test device and the debiased sensor DMOS with reduced heat generation. The effect of the temperature overestimation on I_{sense} is also shown in Fig. 10 (bottom), in which we turned OFF self-heating in the sensor DMOS and set its temperature to that obtained from global model (unidirectional thermo-electric coupling). The reciprocal action of the decreased temperature on the bias point of the sensor transistor is therefore ignored. This situation resembles a *Spice*-like simulation fed by the temperature distribution of a global coupled electro-thermal simulation (e.g., finite volume). Due to the operation close to the *threshold voltage*, this small temperature mismatch causes an overestimation of the sensor current. This exemplifies situations in which one needs to include a fully coupled model because conventional approximation techniques are limiting the accuracy of the result.

The decrease in the sensor current I_{sense} with time in Fig. 10 is caused by the placement of the sensor. As shown in Fig. 9(b), the sensor is located close to the DMOS boundary, away from the temperature hot spot. Due to the operation below the TCP, the current-crowding effect causes increasing temperature difference between the DMOS hot spot and the sensor temperature (middle plot in Fig. 11). As the temperature difference widens, the current ratio shifts in favor of the main DMOS, resulting in a lower current in the sensor branch. This shift can be observed for all three values of R_{shunt}, as shown by means of the current ratio $k = I_D/I_{sense}$ in Fig. 11. By comparison, the hypothetical case of isothermal operation (i.e., no self-heating, $T \stackrel{!}{=} T_{amb} = 25\ °C$), the current ratio k remains constant in time (dashed line in Fig. 11, bottom). The match in k and its temporal evolution for three different operating points are a good confirmation of the electrical and thermal performance of the nonmatching grid. In Fig. 12, the change in current ratio is visualized on top of the transistor transconductance map, where the paths describe how the bias points move as a function of time. The sensor path runs parallel to the contour lines of the transconductance, therefore its conductivity does not change. The path of the main DMOS, however, diverges into a region of higher conductance (towards run-away). Since the total current is fixed in the setup, the measured sensor current must decrease.

V. Conclusion

We have utilized advanced FE methodologies to resolve a small, feature rich region inside a larger power DMOS and allowed these regions to interact in a coupled electro-thermal simulation. The novel aspects of the approach are 1) the effective inclusion of a TCAD transistor model in the FE program as a special nonlinear region and 2) the detailed resolution of a µm-sized structure in the large power DMOS by means of the *Nitsche* nonmatching grid technique. Good agreement is found for the experimentally accessible test bed in a critical power transistor operation regime (self-heating below the TCP). The *Nitsche*-type mortaring technique for coupling independently meshed subdomains at nonmatching interfaces has proved to be very robust in this highly nonlinear setting, even for large jumps in the material parameter (e.g., $\Delta\gamma = 10^{10}$ for oxide-metal interfaces).

After the validation of the global model in Section IV-A, we showed an example of a fully coupled electro-thermal analysis with an integrated current sensor resolved. In this case, traditional approximations result in an overestimation of simulated

sensor current. Our methodology on the other hand, is capable of accurately representing very local temperature inhomogeneities or debiasing effects on a small scale in a fully coupled setting. Availability of tools to approach this set of problems also allows us to investigate the location or topology of the sensor, answering questions arising in device/technology design and design optimization. We plan to combine our shape and optimization capabilities (see, e.g., [31]) with the presented numerical simulation scheme for device structure optimization. Subsequent mechanical simulations, in which device degradation (e.g., metal fatigue) requires highly resolved temperature data, are a case in point. Thermally induced strain is a major concern, particularly for sensible regions like sensing regions.

ACKNOWLEDGMENT

The authors would like to thank our coworkers R. Sleik, G. Glatte, and B. Steinwender for help on the experiment setup, M. Glavanovics for fruitful discussions, and S. Decker for proving the TCAD model. S. de Fillipis original work on modeling the device and current-controlled simulations is greatly acknowledged. A. Hüppe implemented the *Nitsche* method.

REFERENCES

[1] M. Nelhiebel, R. Illing, C. Schreiber, S. Wöhlert, S. Lanzerstorfer, M. Ladurner, C. Kadow, S. Decker, D. Dibra, H. Unterwalcher, M. Rogalli, W. Robl, T. Herzig, M. Poschgan, M. Inselsbacher, M. Glavanovics, and S. Fraiss, "A reliable technology concept for active power cycling to extreme temperatures," *Microelectron. Rel.*, vol. 51, no. 9, pp. 1927–1932, 2011.

[2] V. Khemka, V. Parthasarathy, R. Zhu, A. Bose, and T. Roggenbauer, "Detection and optimization of temperature distribution across large-area power MOSFETs to improve energy capability," *IEEE Trans. Electron Devices*, vol. 51, no. 6, pp. 1025–1032, Jun. 2004.

[3] P. Alpern, P. Nelle, E. Barti, H. Gunther, A. Kessler, R. Tilgner, and M. Stecher, "On the way to zero defect of plastic-encapsulated electronic power devices—Part I: Metallization," *IEEE Trans. Device Mater. Rel.*, vol. 9, no. 2, pp. 269–278, Jun. 2009.

[4] P. Alpern, P. Nelle, E. Barti, H. Gunther, A. Kessler, R. Tilgner, and M. Stecher, "On the way to zero defect of plastic-encapsulated electronic power devices—Part II: Molding compound," *IEEE Trans. Device Mater. Rel.*, vol. 9, no. 2, pp. 279–287, Jun. 2009.

[5] T. Smorodin, J. Wilde, P. Alpern, and M. Stecher, "A temperature-gradient-induced failure mechanism in metallization under fast thermal cycling," *IEEE Trans. Device Mater. Rel.*, vol. 8, no. 3, pp. 590–599, Sep. 2008.

[6] W. Kanert, "Active cycling reliability of power devices: Expectations and limitations," *Microelectron. Rel.*, vol. 52, no. 910, pp. 2336–2341, 2012.

[7] M. Nelhiebel, R. Illing, T. Detzel, S. Wöhlert, B. Auer, S. Lanzerstorfer, M. Rogalli, W. Robl, S. Decker, J. Fugger, and M. Ladurner, "Effective and reliable heat management for power devices exposed to cyclic short overload pulses," *Microelectron. Rel.*, vol. 53, nos. 9–11, pp. 1745–1749, 2013.

[8] T. Veijola and M. Andersson, "Combined electrical and thermal parameter extraction for transistor model," in *Proc. Eur. Conf. Circuit Theory Design*, 1997, pp. 754–759.

[9] P. Bagnoli, C. Casarosa, M. Ciampi, and E. Dallago, "Thermal resistance analysis by induced transient (TRAIT) method for power electronic devices thermal characterization. I. Fundamentals and theory," *IEEE Trans. Power Electron.*, vol. 13, no. 6, pp. 1208–1219, Nov. 1998.

[10] Y. Gerstenmaier, A. Castellazzi, and G. Wachutka, "Electrothermal simulation of multichip-modules with novel transient thermal model and time-dependent boundary conditions," *IEEE Trans. Power Electron.*, vol. 21, no. 1, pp. 45–55, Jan. 2006.

[11] A. Castellazzi, "Comprehensive compact models for the circuit simulation of multichip power modules," *IEEE Trans. Power Electron.*, vol. 25, no. 5, pp. 1251–1264, May 2010.

[12] P. Evans, A. Castellazzi, and C. Johnson, "Automated fast extraction of compact thermal models for power electronic modules," *IEEE Trans. Power Electron.*, vol. 28, no. 10, pp. 4791–4802, Oct. 2013.

[13] M. Pfost, C. Boianceanu, H. Lohmeyer, and M. Stecher, "Electrothermal simulation of self-heating in DMOS transistors up to thermal runaway," *IEEE Trans. Electron. Devices*, vol. 60, no. 2, pp. 699–707, Feb. 2013.

[14] H. Köck, S. de Filippis, M. Nelhiebel, M. Glavanovics, and M. Kaltenbacher, "Multiscale FE modeling concepts applied to microelectronic device simulations," in *Proc. IEEE Int. Conf. Thermal, Mech. Multi-Phys. Simul. Exp. Microelectron. Microsyst.*, 2013, pp. 1–5.

[15] S. Eiser, M. Kaltenbacher, and M. Nelhiebel, "Non-conforming meshes in multi-scale thermo-mechanical finite element analysis of semiconductor power devices," in *Proc. 14th Int. Conf. Thermal, Mech. Multi-Phys. Simul. Exp. Microelectron. Microsyst.*, Apr. 2013, pp. 1–7.

[16] H. Köck, S. Eiser, and M. Kaltenbacher, "Electrothermal multiscale modeling and simulation concepts for power electronics," *IEEE Trans. Power Electron.*, vol. 31, no. 4, pp. 3128–3140, Apr. 2015.

[17] M. Kaltenbacher, *Numerical Simulation of Mechatronic Sensors and Actuators: Finite Elements for Computational Multiphysics*, 3rd ed. Berlin, Germany: Springer, 2015.

[18] T. J. R. Hughes, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Mineola, NY, USA: Dover, Aug. 2000.

[19] R. W. Lewis, K. Morgan, H. R. Thomas, and K. N. Seetharamu, *The Finite Element Method in Heat Transfer Analysis*. New York, NY, USA: Wiley, 1996.

[20] C. Bernardi, Y. Maday, and A. T. Patera, "Domain decomposition by the mortar element method," in *Asymptotic and Numerical Methods for Partial Differential Equations With Critical Parameters*. New York, NY, USA: Springer, 1993, pp. 269–286.

[21] B. Wohlmuth, "A mortar finite element method using dual spaces for the lagrange multiplier," *SIAM J. Numer. Anal.*, vol. 38, pp. 989–1012, 2001.

[22] M. Kaltenbacher, "Advanced simulation tool for the design of sensors and actuators," *Procedia Eng.*, vol. 5, pp. 597–600, 2010.

[23] A. Popp, M. Gitterle, M. W. Gee, and W. A. Wall, "A dual mortar approach for 3d finite deformation contact with consistent linearization," *Int. J. Numer. Methods Eng.*, vol. 83, no. 11, pp. 1428–1465, 2010.

[24] P. Hansbo and J. Hermansson, "Nitsche's method for coupling non-matching meshes in fluid-structure vibration problems," *Comput. Mech.*, vol. 32, no. 1, pp. 134–139, 2003.

[25] S. Sitzmann, K. Willner, and B. I. Wohlmuth, "A dual lagrange method for contact problems with regularized frictional contact conditions: Modelling micro slip," *Comput. Methods Appl. Mech. Eng.*, vol. 285, pp. 468–487, 2015.

[26] N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics: Converters, Applications, and Design*, B. Zobrist, Ed. New York, NY, USA: Wiley, 2003.

[27] D. Dibra, M. Stecher, S. Decker, A. Lindemann, J. Lutz, and C. Kadow, "On the origin of thermal runaway in a trench power Mosfet," *IEEE Trans. Electron. Devices*, vol. 58, no. 10, pp. 3477–3484, Oct. 2011.

[28] S. de Filippis, V. Kosel, D. Dibra, S. Decker, H. Köck, and A. Irace, "ANSYS based 3D electro-thermal simulations for the evaluation of power MOSFETs robustness," *Microelectron. Rel.*, vol. 51, no. 11, pp. 1954–1958, 2011.

[29] S. de Filippis, R. Illing, M. Nelhiebel, S. Decker, H. Kock, and A. Irace, "Validated electro-thermal simulations of two different power Mosfet technologies and implications on their robustness," in *Proc. 25th Int. Symp. Power Semiconductor Devices ICs*, May. 2013, pp. 325–328.

[30] J. Sauveplane, P. Tounsi, E. Scheid, and A. Deram, "3d electro-thermal investigations for reliability of ultra low ON state resistance power MOSFET," *Microelectron. Rel.*, vol. 48, nos. 8/9, pp. 1464–1467, 2008.

[31] F. Wein, M. Kaltenbacher, and M. Stingl, "Topology optimization of a cantilevered piezoelectric energy harvester using stress norm constraints," *Struct. Multidisciplinary Optimiz.*, vol. 48, no. 1, pp. 173–187, 2013.

Sebastian Eiser studied at Technische Universität Berlin, Berlin, Germany and Duke University, Durham, NC, USA. He received the Diploma degree in theoretical physics from Technische Universit at Berlin, Germany.

He was with the European Space Agency, Frascati, Italy, where he learnt about spaceborne measurement principles, earth observing satellites, Pecorino, and Italian. During the Ph.D. degree, he enjoyed learning techniques and methods from the broad palette of skills needed in computational engineering. As a procrastination master, he certainly overdid tinkering with the Linux shell and TikZ/LaTeX graphics, but eventually managed to publish this paper.

Mirko Bernardoni received the M.Sc. degree in electronic engineering, in 2009, and the Ph.D. degree in information technologies, in 2012, both from University of Parma, Parma, Italy.

In 2012, he was a Post-doctoral Researcher at the Centre for Device Technology and Reliability, University of Bristol, Bristol, U.K., and since 2013, is employed as a Research and Development Engineer for simulation techniques at the Kompetenzzentrum für Automobil und Industrieelektronik, Villach, Austria. His research activities include thermal and electro-thermal modeling of automotive devices with finite element method, and more generally on the field of reliability and degradation analysis of power electron devices.

Michael Nelhiebel received the M.Sc. degree in physics from the Vienna University of Technology, Vienna, Austria, and the Ph.D. degree in solid-state physics from Ecole Centrale Paris, Paris, France, working on interferometry in electron energy loss spectrometry.

In 1999, he joined Infineon Technologies, Austria, as a Reliability Engineer of the silicon wafer production. He is currently a Senior Staff Engineer with the Department of the Quality, Automotive Business Division, responsible during the development phase for technology related product reliability. He has coordinated the qualification of major automotive technology platforms and participates in research activities of Infineon Technologies Austria targeting technology reliability.

Manfred Kaltenbacher (M'93) received the Dipl.-Ing. degree in electrical engineering from the Technical University of Graz, Graz, Austria, in 1992, the Ph.D. degree in technical science from the Johannes Kepler University of Linz, Linz, Austria, in 1996, and the Habilitation degree from Friedrich-Alexander-University, Erlangen-Nuremberg, Germany, in 2004.

In 2008, he became a Full Professor for applied mechatronics at Alps-Adriatic University Klagenfurt, Klagenfurt, Austria. In 2012, he moved to Vienna University of Technology, Austria, as a Full Professor for measurement and actuator technology. He is the author and coauthor of a book, five book chapters, and more than 60 peer-reviewed journal publications. His research interests include computer-aided engineering of sensors and actuators with special emphasis on numerical simulation techniques.

Dr. Kaltenbacher is a Member of the IEEE Society, the AIAA Society, the German Society of Electrical Engineers (VDE), the German Society of Acoustics (DEGA), the Austrian Acoustic Association, and the International Compumag Society.