A new method for prediction of 2D chip temperature distributions in general purpose drive load profiles

Alexander Philippou, Jan P. Meissner, David Übelacker, Christian R. Müller,
Franz-Josef Niedernostheide

Infineon Technologies AG, Germany

Corresponding author: Alexander Philippou, alexander.philippou@infineon.com

Abstract

Using parallel computing and a divide-and-conquer technique, a fast and accurate method is presented to enable 2D power-loss and temperature distributions of chips for load profiles in pulse-width modulated (PWM) two-level inverter applications. Because variations of the temperature and power losses on the chips are taken into account, a more accurate determination of the maximum module current for application-specific load and overload profiles is possible.

1 Introduction

The maximum junction temperature in an IGBT device is an important parameter for the design of the inverter and cooling system. Especially for larger chips, e.g. for inverter output power beyond 10 kW, the level of temperature inhomogeneity on the chip increases. With the trend toward higher virtual junction temperatures of 175°C, the hottest regions on the chip will be even higher. Therefore, it becomes increasingly important to assess the temperature inhomogeneity on the chip in an application-specific load profile. Several approaches exist to describe the electrical losses and thermal distribution of a power device in a two-level pulse-width modulated (PWM) inverter application [1-3]. Depending on the applied method, approximations are made about the complexity of the thermal or electrical model of either inverter or semiconductor device. The most rigorous approach to the thermal model is a full-scale finite element simulation, in combination with an elaborate SPICE model for the semiconductor electrical losses. For a single cycle of the PWM, computation times are long. While the thermal model complexity can be drastically reduced by applying model order reduction [4], accurate simulations of full load profiles consisting of thousands of cycles still remain difficult. We propose a new method based on a temperature-dependent loss model for the semiconductor device, in combination with a divide-and-conquer technique for the thermal network to calculate the temperature distribution on the chip. Its main advantage is a considerably faster computation time while maintaining important temperature distribution information. At the same time, the method allows for a high degree of parallelization. This allows for simulations of full load profiles in time ranges of minutes and faster.

2 Method overview

The IGBT and diode are divided into segments, where thermal-impedance coupling $Z_{th,m->n}$ between a powered segment *m* and its resulting temperature coupling into segment *n* is described by a foster network. A 3D FEM model, consisting of a Cu-Al2O3-Cu DCB substrate with six diodes and six IGBTs, is used to calculate the parameters

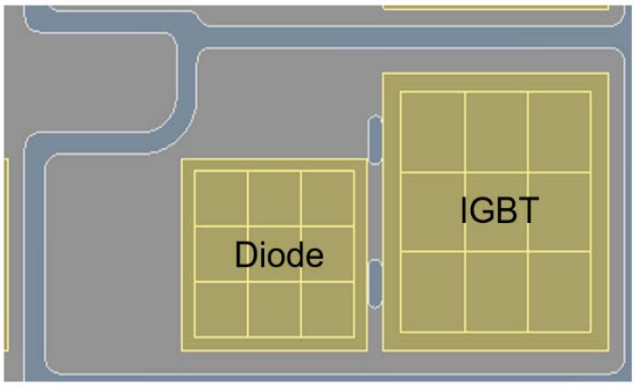

Fig. 1: Module layout for the segmented chips is shown for a 1200 V, 75 A IGBT and diode.

ISBN 978-3-8007-5245-4

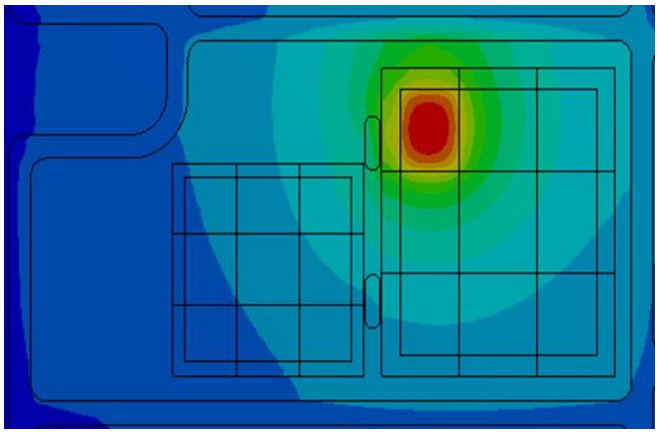

Fig. 2: Temperature distribution for the heating of one chip segment as resulting from a FEM calculation.

for this network. The substrate is connected to a heatsink with thermal grease. One diode and one IGBT are divided into 9 segments (**Fig. 1**). By using different automation scripts, 28 transient simulations are performed. To obtain the cross coupling of a segment with another segment, each chip segment is heated separately. **Fig. 2** shows the temperature distribution when only one segment is heated up.

This simulation data is used to build up a 28x28 $Z_{th}(t)$ matrix which includes all cross coupling between the chips and the segments.

Starting from this matrix, the Foster parameters consisting in sum of 9408 thermal resistance coefficients $R_{th,i}$ and time constants τ_i, are calculated. As a result, a six node Foster network was used for each coupling element to describe the corresponding thermal impedance coupling between elements. As reference, a simulation with homogeneous heating over all segments is used, and the thermal resistance is calculated. The temperature distribution is shown in **Fig. 3**. This

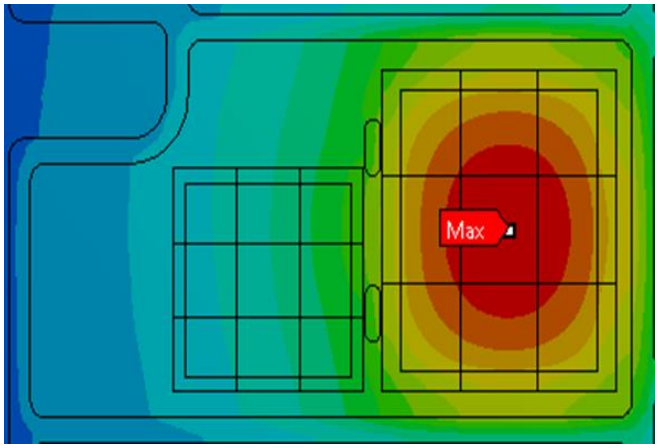

Fig. 3: Temperature distribution for heating of all chip segments as resulting from a FEM calculation.

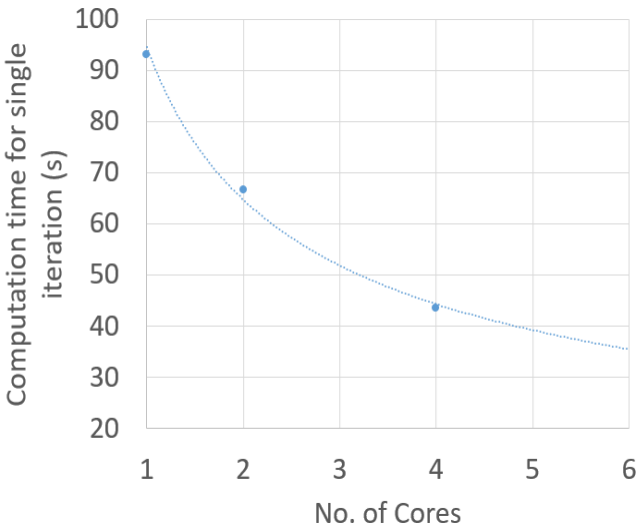

Fig. 4: Computation (wall clock) time for a full load profile with 15000 cycles for a single iteration is shown as evaluated on a 2.7 GHz processor.

value can be used as reference for the thermal resistance calculated from the segmented simulations. By summing up and weighting the thermal resistances of the segments with regard to their active area, an overall thermal resistance can be calculated. In this case, the difference between both thermal resistances is smaller than 0.5%. Once the Z_{th} matrix elements are calculated, they can be fed into the algorithm proposed in [5], where the coupling of elements considered was restricted to the IGBT and diode elements of neighbored chips. The equation for temperature T_n of segment *n* is then given by

$$T_n(\mathrm{t}) = \sum_m P_m(\mathrm{t}, \mathrm{T}_m) * \frac{dZ_{th,m\to n}(t)}{dt} \qquad (1)$$

Here, $P_m(t,T_n)$ denotes the time and temperature-dependent losses of either IGBT or diode in segment m, and $Z_{th,m->n}$ the corresponding thermal impedance coupling from segment m to n. This sum can easily be distributed to multiple processes to speed up calculation, and can be solved for each segment separately. Once a solution is obtained, the new temperature in each segment is used as a starting point for the next iteration, until convergence is reached, similar as in [5]. Thereby convergence is reached within a couple of iterations. As the algorithm is also suited for full load profile computations, its computation time as shown in **Fig. 4** is applied to a 300 s load profile and an output frequency of 50 Hz, therefore calculating 15000 cycles. The scalability of the algorithm with an increasing number of processor

ISBN 978-3-8007-5245-4

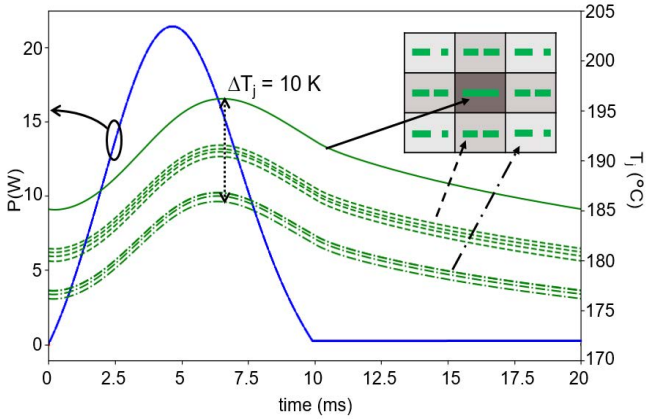

Fig. 5: The different temperature and power losses of each segment are shown, where dashed-dotted lines represent chip corners, dashed lines edges, and solid lines the chip center. Blue represents the power losses, shown on the left axis, while green represents the junction temperatures, shown on the right axis.

cores can be observed where an increase in cores by a factor of 4 reduces computation time by a factor of 2. It eventually saturates due to overhead computation costs of 5 s for each segment. Therefore, the further increase of cores is less effective for 9 segments.

3 Results

3.1 Constant load profiles

A constant load profile of a sinusoidal RMS current of I_{rms} = 48 A, modulation frequency f_{mod} = 2.7 kHz, output frequency f_{out}=50 Hz, ambient temperature T_{amb} = 50 °C, thermal heatsink resistance R_{thha} = 1.2 K/W, and phase shift cos(Φ) = 0.85, was simulated. In a first step, the temperature dependence of the power losses was eliminated in the model. The results for the different segments are shown in **Fig. 5**. Clearly, the hottest region is in the chip center, while the coolest regions are in the corners. Temperature spreading from corner to center is $\Delta T_{j,max}$ = 10 K. Taking into account the temperature dependence of the losses P(t,T), a further increase of the temperature spreading to $\Delta T_{j,max}$ = 14 K can be observed as shown in **Fig. 6**. Values observed in the experiment are in the range of up to $\Delta T_{j,max}$ = 40 K [6]. The difference can be attributed to two effects. In [6], a larger chip was used compared to the chip size in this paper, which increases temperature inhomogeneity. Additionally, in the simulation approach here, only three segments along a chip dimension were evaluated. Increasing the number of segments is in principle possible and will improve accuracy. As

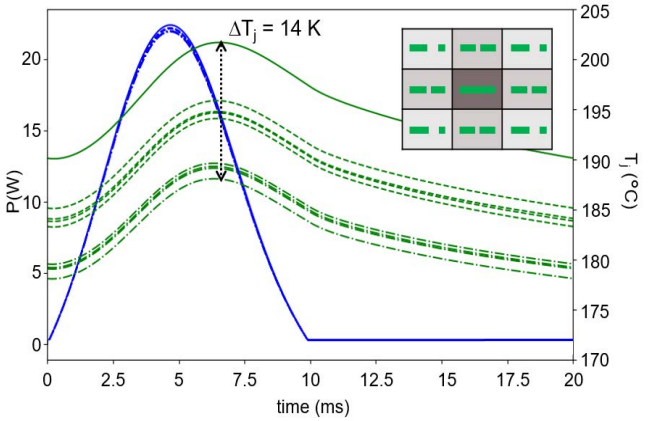

Fig. 6: Temperature distribution for different segments with temperature-dependent losses for each segment shown for I_{rms} = 48 A.

a demonstration example, we have chosen 3x3 segments to limit computational complexity, especially when we later consider full load profiles. Using this approach, calculations can run with multiple processes, but on a single machine.

3.2 Module layout effects

In a next step, the method was applied to different module layouts to investigate the effect of temperature spreading when chips are more widely spaced. The results are shown in **Fig. 7**,

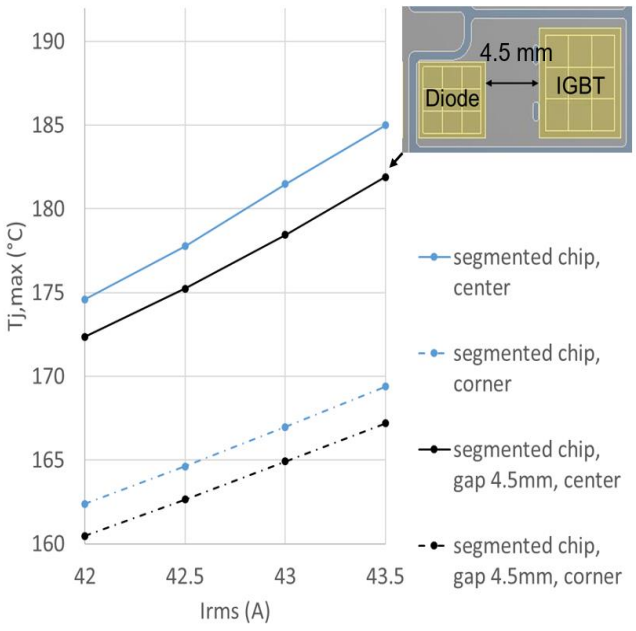

Fig. 7: Periodic $T_{j,max}$ over I_{rms} current is shown. Increased spacing between chips to 4.5 mm reduces $T_{j,max}$ by 5 K, with a similar temperature difference between chip center and corner for both spacings.

ISBN 978-3-8007-5245-4

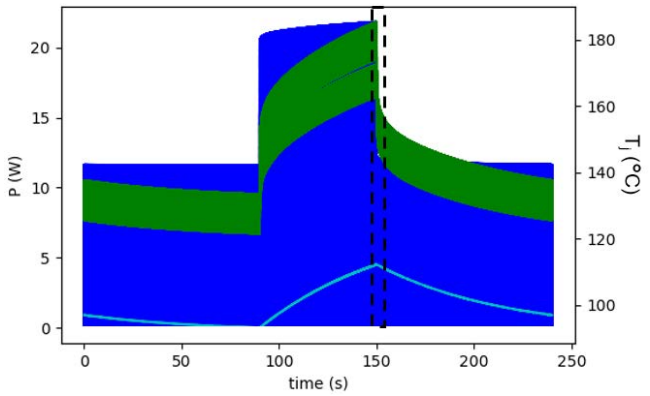

Fig. 8: Overview of a periodic load profile where I_{rms} = 32 A for 180 s, and 48 A for 60 s. Temperature difference between the center region of the chip upper green line), compared to the corner region (lower green line) can be observed. The solid black lines depict a region for zooming into the overload pulse. The cyan colored line shows the temperature of the heatsink.

where a comparison was made between the layout shown in **Fig. 2** and an increased gap of 4.5 mm between IGBT and diode. As can be seen, the maximum junction temperature is reduced for increased spacing, while the temperature spread on the chip remains nearly unaffected. This effect is especially important if chip shrinkage effects as part of new IGBT generations are investigated. As generally a smaller chip size increases R_{th} from junction to case, the increased spacing between the chips partly counteracts this effect.

3.3 GPD load profiles

Finally, also a typical general purpose drive (GPD) load profile [7] was studied where the temperature of each chip segment for both IGBT and diode was evaluated. For the load profile, a current of I_{rms} = 32 A was applied for 180 s with a sudden increase to 48 A for 60 s. The heatsink time constant was

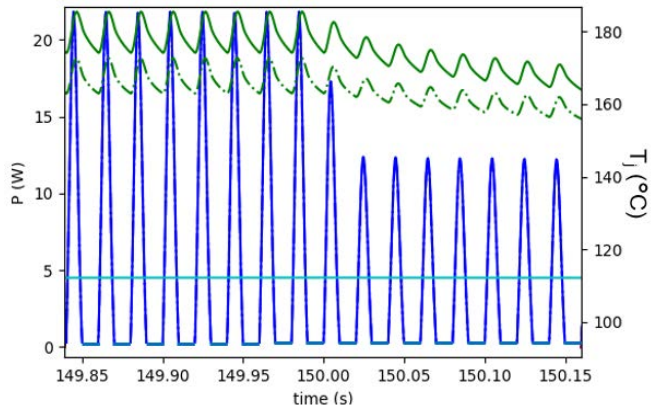

Fig. 9: Zoom into region marked by the dashed black lines in **Fig. 8**.

selected to be 60 s. The results are shown in **Fig. 8**, where only temperatures of the center and corner chip regions were plotted. A zoom into the dashed region is displayed in **Fig. 9**. To further analyze the effect of temperature inhomogeneity during the load profile, a measure for inhomogeneity, defined as the temperature difference between corner and center chip region, of two different load profiles (**Fig. 10**), was evaluated and plotted over time. The results in **Fig. 11** show in case of profile "Ovl1" a maximum temperature inhomogeneity of 13 K, while for the second load profile "Ovl2" a larger temperature spread of 17 K can be observed. It occurs in case of "Ovl1" at the end of the overload pulse at t = 150 s (dashed line). In case of "Ovl2" it occurs during a short 3 s pulse to I_{rms} = 60 A before going into a 60 s I_{rms} = 42 A overload pulse. It should be noted that care was taken to ensure that the same maximum junction temperature was reached for both overload profiles.

4 Conclusion

We demonstrated a new method for computation of temperature distributions in load profiles for a 2-level PWM inverter application. The method divides both IGBT and diode into segments, where the coupling between each segment is described by a Foster network. The new method allows for computation times in the range of seconds to minutes for full load profiles, while considering temperature-dependent electrical losses in each segment. Several applications have been demonstrated. The impact of temperature

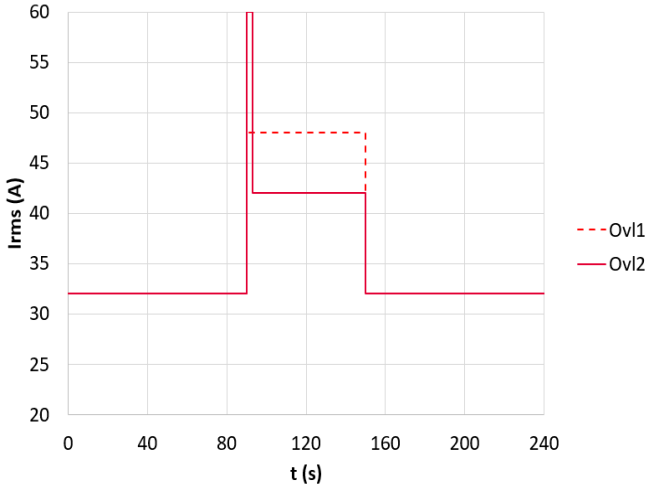

Fig. 10: Investigated load profiles labeled "Ovl1" and "Ovl2." Both current values of the load profiles were adjusted such that the device reaches same $T_{j,max}$. Load profile "Ovl2" shows a stronger variation in load current.

ISBN 978-3-8007-5245-4

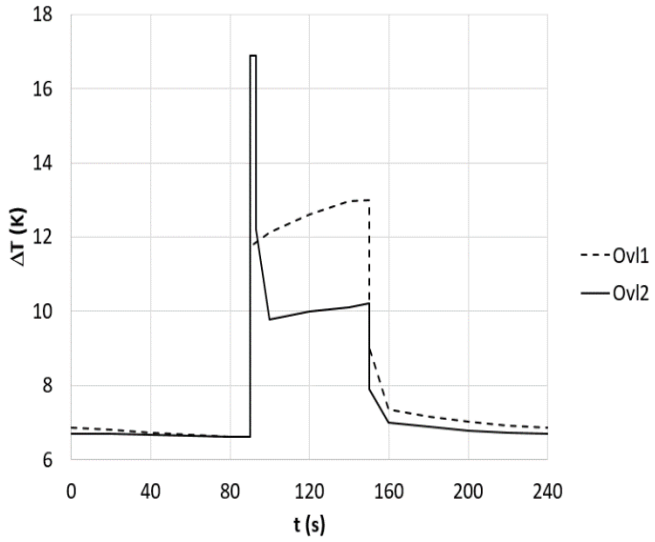

Fig. 11: Temperature difference ΔT between center and corner segment of the chip are plotted over time for the two different load profiles labeled “Ovl1” and “Ovl2”.

dependence of electrical losses was investigated. These contributed roughly by 30% to the temperature inhomogeneity, while the remaining percentage was due to the Z_{th} coupling between segments. A module layout study showed that a larger separation between chips can reduce the maximum junction temperatures, while its effects on temperature inhomogeneity reduction on the chip is negligible. Furthermore, different load profiles were analyzed showing that the largest temperature inhomogeneity on the chip occurs during short-term scale overcurrent pulses compared to the heatsink time constant.

5 References:

[1] H. A. Mantooth and A. R. Hefner, "Electro-thermal simulation of an IGBT PWM inverter", Proceedings of IEEE Power Electronics Specialist Conference - PESC '93, 1993, pp. 75-84

[2] H. Huang, A. T. Bryant and P. A. Mawby, "Electro-thermal modelling of three phase inverter", Proceedings of the 2011 14th European Conference on Power Electronics and Applications, 2011, pp. 1-7

[3] R. Bornoff, A. Vass-Varnai, B. Blackmore, G. Wang and V. H. Wong, "Full-circuit 3D electro-thermal modeling of an IGBT Power Inverter", 33rd Thermal Measurement, Modeling & Management Symposium (SEMI-THERM), 2017, pp. 29-35

[4] T. Bechtold, E. B. Rudnyi and J. G. Korvink, “Dynamic electro-thermal simulation of microsystems - a review”, J. Micromech. Microeng.**15,** R17, 2005

[5] A. Philippou, C. R. Müller, F.-J. Niedernostheide, B. Sahan, “Accurate Prediction of Thermal Runaway at Blocking Conditions in Drives Applications”, Proc. PCIM Europe, 2019, Germany, pp. 91-95

[6] U. Scheuermann, R. Schmidt, “Investigations on the VCE(T)-Method to Determine the Junction Temperature by Using the Chip Itself as Sensor”, Proc. PCIM Europe, 2009, Germany, pp. 802-807

[7] S. Buschhorn et al., “Impact of load profiles on power module design – a detailed analysis based on 7th generation of IGBT and diode technology”, Proc. PCIM Europe, 2018, Germany, pp. 570-577