Numerical investigation on thermal crosstalk of silicon dies in high voltage IGBT modules

Thomas Hunger und Oliver Schilling
Infineon Technologies AG, Max-Planck-Straße 5, D-59581 Warstein

Abstract

The paper investigates numerically the mutual thermal interaction of the dies within a high power IGBT module. The influence of the cooling environment on the coupling has been studied in detail. Thermal compact models based on RC networks were deduced from the thermal impedance curves. They allow for the study of the importance of the crosstalk during inverter operation.

1 Introduction

The knowledge of the virtual junction temperature (T_j) of power semiconductor devices under operating conditions is a vital part in the design phase of power applications regarding overload and reliability. Especially, semiconductor manufacturers specify a maximum allowed junction temperature for device operation which may no be exceeded under every operating conditions.
High-voltage (HV) IGBT modules usually comprise a parallel connection of many of two different die types, namely the IGBTs itself and the anti-parallel diodes acting during freewheeling (FWD). Once one die is actively heated the nearby dies will also show a temperature rise. Simplifying approaches for determining T_j during operation either neglect that cross coupling or use some heuristic approaches.
In the near past there have been many publications dealing with the description of thermal crosstalk of multi-chip assemblies in IGBT modules [1-5]. Based on finite element approaches compact models had been deduced and consequences of crosstalk-heating during inverter operation were discussed.
In our paper we use the finite element method to calculate the thermal impedance curves which are sufficient for the full description of the whole thermal system. The influence of the cooling environment on the coupling strength is then discussed in more detail. Due to the structure of many parallel acting dies in the HV-IGBT package under study a simplified thermal model is proposed. Compact models are deduced on that basis which finally allow for the study of the temperature swing at different operation points.

2 Single die approach

2.1 Model description

The device under study is a standard module used for traction application with a footprint of 190x140 mm² (fig. 1).

It contains 24 paralleled IGBT and 12 paralleled diode dies equally distributed on 6 ceramic substrates.

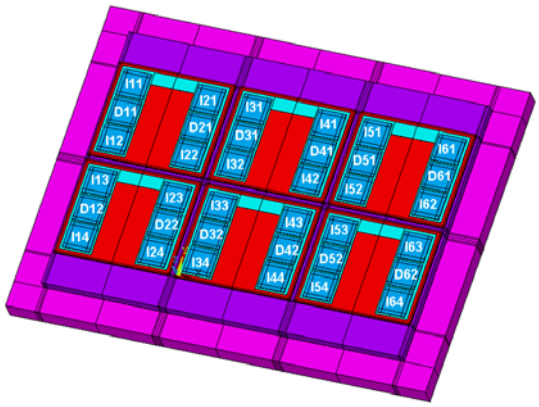

Fig. 1 The HV-IGBT module under consideration. The positions of the IGBTs (“I”) and diode (“D”) dies are marked. The module is mounted to a heat sink plate.

The base plate is made out of AlSiC, the ceramic is AlN. Solder layers connect the dies to the ceramic substrates and the substrates to the base plate, respectively. The cross section showing all layers is shown in fig. 2.

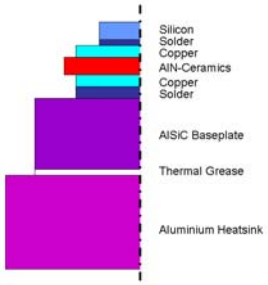

Fig. 2 Schematic cross section of the IGBT module (not to scale)

The thermal simulations were performed using the commercial finite element tool Ansys [6]. Temperature dependent material data were used. The heat sink was modelled as an aluminium

block connected via a thermal grease layer to the power module. On the back side of the heat sink convective boundary conditions characterized by a heat transfer coefficient α and an ambient temperature T_a were applied. This effective description of a cooler is necessary to reduce the amount of computation resources. A full fluid dynamic treatment taking into account all details of the inner structure of the cooler is hence no longer needed [2].
The experimentally measured heat sink thermal resistance $R_{th,ha}$ can be analytically related to α

$$R_{th,ha} = d_{\mathrm{Al}} \Big/ (\lambda_{\mathrm{Al}} \cdot A_{\mathrm{Module}}) + 1 \Big/ (\alpha \cdot A_{\mathrm{Cooler}}) \cdot \quad (1)$$

Fig. 3 shows $R_{th,ha}$ as a function of the heat transfer coefficient α obtained using different methods. Additionally, typical R_{th} regimes for forced air as well as water coolers are indicated. It is again clearly visible that the effective heat sink description by α is a suitable approach.

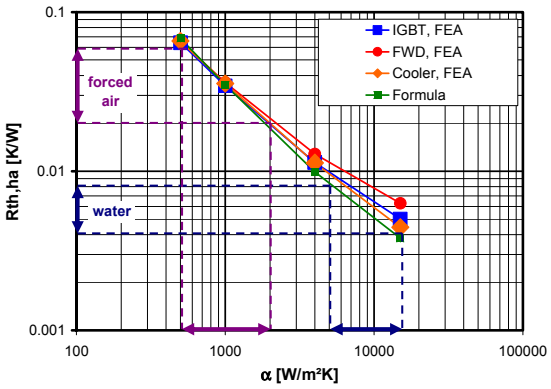

Fig. 3 $R_{th,ha}$ as a function of the heat transfer coefficient applied to the backside of the aluminium cooling plate. The curves were obtained via heating of all IGBT ("IGBT, FEA") and all FWD dies ("FWD, FEA"). Additionally, a homogeneous area heat load on the module's footprint size was applied ("Cooler, FEA"). The analytical result from formula eq. 1 is also given.

2.2 Thermal impedance and junction temperature

The temperature evolution in a thermal system caused by power injection can be described by the thermal impedance formalism. The system under study consists of 36 spatially separated heat sources (=dies) and one heat sink. The power loss in the dies is modelled via a homogeneous heat density applied only to the volume of the projected active area, i.e. no heat is present in the high voltage termination at the dies' edge.
During the application of the step-like heat pulse the temperatures were recorded at the "position" of the junction, the base plate and the top of the heat sink below each die. The usual definition of the thermal impedance with a step like constant power P_Y in die Y and a temperature rise in die X is written as

$$Z_{th,XY} = \frac{T_{j,X} - T_a}{P_Y} \quad (2)$$

For all 36 dies and arbitrary heat profile in each die one obtains with respect to the ambient temperature

$$\begin{pmatrix} T_j^{I11}(t) \\ \vdots \\ T_j^{I64}(t) \\ T_j^{D11}(t) \\ \vdots \\ T_j^{D62}(t) \end{pmatrix} = \begin{pmatrix} \hat{Z}_{I11} & \cdots & \hat{Z}_{I11-I64} & \hat{Z}_{I11-D11} & \cdots & \hat{Z}_{I11-D62} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ \hat{Z}_{I62-I11} & \cdots & \hat{Z}_{I64} & \hat{Z}_{I64-D11} & \cdots & \hat{Z}_{I64-D62} \\ \hat{Z}_{D11-I11} & \cdots & \hat{Z}_{D11-I64} & \hat{Z}_{D11} & \cdots & \hat{Z}_{D11-D62} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \hat{Z}_{D62-I11} & \cdots & \hat{Z}_{D62-I64} & \hat{Z}_{D62-D11} & \cdots & \hat{Z}_{D62} \end{pmatrix} \begin{pmatrix} P_{I11}(t) \\ \vdots \\ P_{I64}(t) \\ P_{D11}(t) \\ \vdots \\ P_{D62}(t) \end{pmatrix} + \begin{pmatrix} T_a \\ \vdots \\ T_a \\ T_a \\ \vdots \\ T_a \end{pmatrix} \quad (3)$$

The impedance matrix entries are convolution operator [7,8] as

$$\hat{Z}_{th} P(t) = \int_{-\infty}^{t} d\tau \cdot \dot{Z}_{th}(t-\tau) \cdot P(\tau) \,. \quad (4)$$

The main diagonal entries are the usual thermal impedance junction to ambient, $Z_{th,ja}$, of each die whereas the off-diagonals describe the couplings.
Throughout this investigation the virtual junction temperature is measured as the maximum temperature at the chip surface for the FWDs and the bottom of the chip for the IGBTs. This procedure takes into account the position of the relevant pn-junction for temperature determination in a usual Z_{th}-measurement setup using the forward voltage drop for temperature determination. However, the forward voltage drop averages the temperature in a certain way over the chip. Therefore, the hot spot used in evaluation of the FEA results tends to overestimate the thermal resistance $R_{th,jc}$.

2.3 Single die heating

For the following calculations a heat transfer coefficient of α=4000W/m²K is used. Fig. 4 shows the thermal impedance as the step response to a constant heat power applied in IGBT I11.

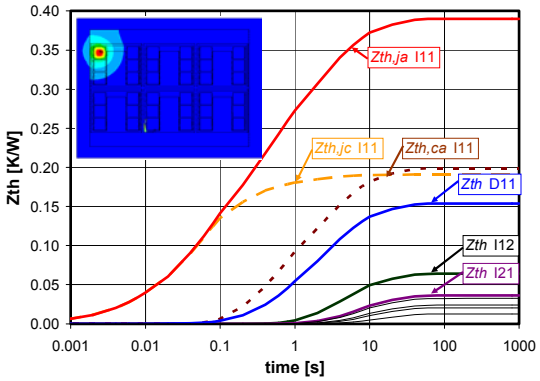

Fig. 4 Thermal response of different dies while applying heat to IGBT I11. The part $Z_{th,jc}$ as well as $Z_{th,ca}$ of I11 are shown for comparison.

The temperature rises in I11 immediately after applying the heat. The temperature increase in all other dies follow with a time delay which is clearly seen in fig. 4 as a zero value of the impedance for these small times. First, the next neighbouring die, diode D11, starts at ~100ms with the passive temperature increase. Later, the temperature rise in I12 and I21 is visible. All other dies follow with a more pronounced delay but with reduced absolute strength of the static coupling value. This behaviour can be understood simply by the propagation of the heat which, beside its vertical propagation down to the heat sink, also shows lateral spreading.
Interestingly, the crosstalk impedances and the direct case-to-ambient impedance (also shown in fig. 4) show similar shape, notably seen for die D11. Both show a time delay before the initial temperature rise after heating the die. Based on that one might think about a fitting of the crosstalk impedance as a scaled function of $Z_{th,ca}$ depending on the die separation and the cooling environment α. However, this seems impossible to perform because of the complexity of the real heat path.

2.4 Influence of heat sink type

The next section considers the influence of the heat sink on the coupling strength. Fig. 6 and fig. 7 show calculations of heating IGBT I32. The heat transfer coefficient applied to the back side of the heat sink were taken as α=500W/m²K (weak forced air cooler) and 15000W/m²K (strong water cooler), respectively.

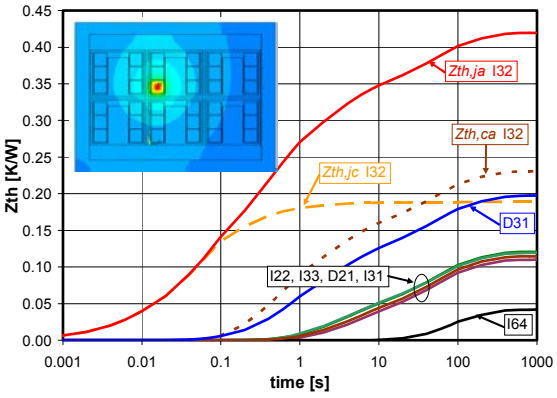

Fig. 5 Thermal response of different dies while applying heat to IGBT I32 at α=500W/m²K. The part $Z_{th,jc}$ as well as $Z_{th,ca}$ for I32 are shown for comparison.

The direct impedance $Z_{th,ja}=Z_{th,jc}+Z_{th,ca}$ of I32 is considered first. For small periods of time <100ms the heat propagates within the module, $Z_{th,ja}=Z_{th,jc}$. Both thermal response curves are equal and independent of the cooling environment. Later (>2s), the heat sink starts governing the heat spreading leading to a larger thermal resistance for the low-α-coolers (cf. insets in figs. 5 & 6).
The coupling between the dies is more pronounced for coolers having low α-values. The beginning of the temperature rises in the non-heated dies is similar whereas the absolute static values as well as the shapes are different. This is again due to the fact of a larger heat spreading in low-α-coolers and clearly visible when tracing the behaviour of the outermost die I64. A significant temperature increase is measured in fig. 5 whereas no reaction is seen in fig. 6.

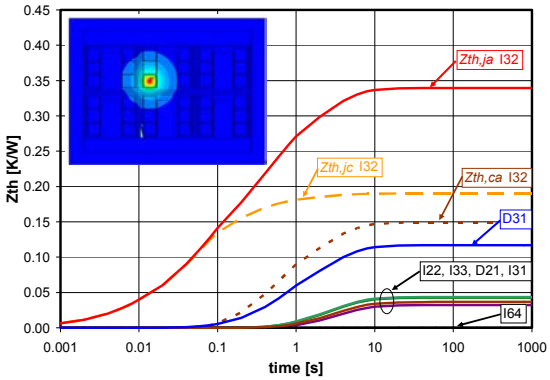

Fig. 6 Thermal response of different dies while applying heat to IGBT I32 at α=15000W/m²K. The part $Z_{th,jc}$ as well as $Z_{th,ca}$ for I32 are shown for comparison.

To conclude this paragraph it is found that the heat sink has a major impact on the coupling strength of the dies. However, the complexity of the single die approach prohibits the handling by hand. Therefore, a simplified approach is proposed in the next paragraph to circumvent these obstacles.

3 Die type approach

3.1 Impedance simplified

During operation usually all IGBTs and all FWDs are heated at the same time, respectively. If one assumes equal heat dissipation in the dies, i.e. neglecting the small differences during switching due to deviations of the parasitic environments, considerable reduction of the problem will be possible as described below.
Formally, one writes

$$P_{FWD}(t) = P_{D11}(t) = \ldots = P_{D62}(t),$$
$$P_{IGBT}(t) = P_{I11}(t) = \ldots = P_{I64}(t). \quad (5)$$

The Z_{th} curves are now calculated while heating all chips of one type simultaneously. Next, a properly chosen average for T_j of each die type based on the temperatures of each individually contributing die is considered. The maximum temperature in the dies was taken for determining the individual Z_{th} curves so far. Because this overestimates the actually measured R_{th} a simple arithmetic average over all Z_{th} curves of one die type seems more reliable than using the hottest only. With this one can reduce eq. 3 to

$$\begin{pmatrix} T_j^{IGBT} \\ T_j^{FWD} \end{pmatrix} = \begin{pmatrix} \hat{Z}_{th,ja}^{IGBT} & \hat{Z}_{th}^{IGBT\leftarrow FWD} \\ \hat{Z}_{th}^{FWD\leftarrow IGBT} & \hat{Z}_{th,ja}^{FWD} \end{pmatrix} \cdot \begin{pmatrix} P_{IGBT}^{module}(t) \\ P_{FWD}^{module}(t) \end{pmatrix} + \begin{pmatrix} T_a \\ T_a \end{pmatrix} \quad (6)$$

The impedance matrix entries are convolution operators as in eq. 4. Note, the matrix is not longer symmetric because of the different numbers of dies of one type. The power P_{IGBT} and P_{FWD} now contain the loss per module.
The inset of fig. 7 shows the temperature distribution in the module for clarity.

3.2 Heat sink dependence of thermal coupling

This section deals with the influence of the cooler performance on the crosstalk in the formulation given previously. Fig. 7 and 8 show the thermal impedance curves for two different cooler types with α=1000W/m²K and 15000W/m²K. Generally, the statements on coupling from paragraph 2.4 can be applied to the current results.

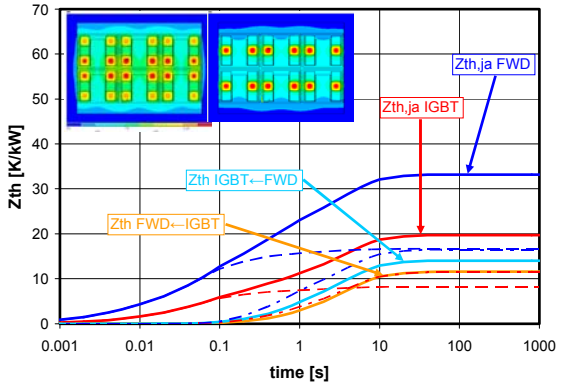

Fig. 7 Thermal response of different die types while applying heat to the IGBTs and FWDs, respectively, for α=15000W/m²K. The crosstalk impedances are also given. The $Z_{th,jc}$ and $Z_{th,ca}$ are shown for comparison (dashed and dash dotted). Left inset: Temperature distribution when IGBTs heated, right inset: FWDs heated.

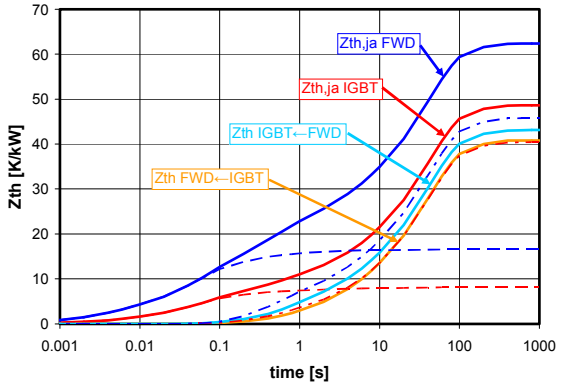

Fig. 8 Thermal response of different die types while applying heat to the IGBTs and FWDs, respectively, for α=1000W/m²K. The crosstalk impedances are also given. The $Z_{th,jc}$ and $Z_{th,ca}$ are shown for comparison (dashed and dash dotted).

The shape of the crosstalk impedances again roughly follows the $Z_{th,ca}$ of the heating chips. There are deviations especially in the rising slope region of the curves.

3.3 Scaling of crosstalk resistance

The results of the previous paragraph suggest a scaling of the thermal coupling impedance with the case-to-ambient values of the heating chip. In fig. 9 the static values are compared for both the IGBT and the FWD. All relevant parts of the thermal resistance are shown. $R_{th,jc}$ remains nearly constant for all α, whereas the cooler $R_{th,ha}$ varies strongly. Accordingly, the crosstalk scales.

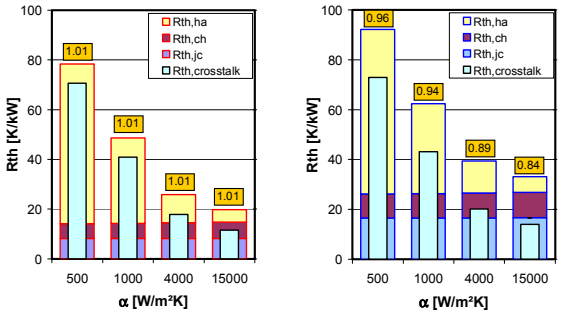

Fig. 9 Thermal resistances as function of the heat transfer coefficient for IGBT (left) and FWD (right). The ratio of the $R_{th,crosstalk}$ and $R_{th,ca}$ of the heating chip is given. The crosstalk is left from IGBT to FWD and right vice versa.

Interestingly, the coupling strength expressed as the ratio of $R_{th,crosstalk}/R_{th,ha}$ is nearly independent on α for the IGBT and shows only a slight variation with α for the FWD. The absolute values, however, varies strongly with α.

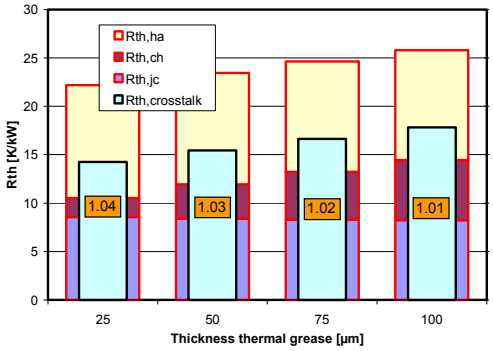

Fig. 10 Thermal resistances for the IGBT shown as function of the thermal grease thickness (α=4000W/m²K).

The study of the crosstalk depending on the thermal grease thickness is shown in fig. 10. The coupling varies again with $R_{th,ca}$ within a range of some percent.
These results suggest that a simple scaling of the $Z_{th,ca}$ to fit the crosstalk impedances seems possible. In the present case this is partly due to the symmetry within the module: every FWD die is surrounded by two IGBT chips. Nevertheless, it might be useful for the user of HV-IGBT modules to have coupling values in the datasheets. That needs more detailed consideration both theoretically and experimentally and should be extended to other layouts. This is beyond the scope of this paper.

4 Compact modelling and load case studies

4.1 Derivation of thermal compact models

So far only the thermal impedance as the reaction of the thermal system to a step power impulse was discussed. The study of inverter operation by means of real loss profiles is of higher interest. The direct calculation of the junction temperature via finite element analysis is not possible due to the overwhelming calculation times required. With the use of the convolution from eq. 4 it is possible to calculate the junction temperature for arbitrary load. Usually, a mapping of Z_{th} to equivalent electrical RC-networks is used. Those compact models can be solved in common SPICE simulators or even analytically.
In our study RC-networks of the Foster type, eq. 7, were used. They allow for an easy fitting of the impedance curves to the resistances and capacitances required. Typically, four terms were used.

$$Z_{th}(t) = \sum R_i \left(1 - \exp\left(-t/\tau_i\right)\right), \quad \tau_i = R_i C_i \quad (7)$$

The fitting describing the mutual thermal interaction need a little more attention because they have a certain time delay as mentioned above. If one accepts negative resistance and capacitance values the description will give suitable results. Fig. 11 shows the comparison between Z_{th} from the FEA and the fits according to eq. 7. The overall match is sufficient. There are little deviations between the FEA and the fitting of the crosstalk impedance below 100ms. However, this is acceptable with respect to the range of required accuracy.

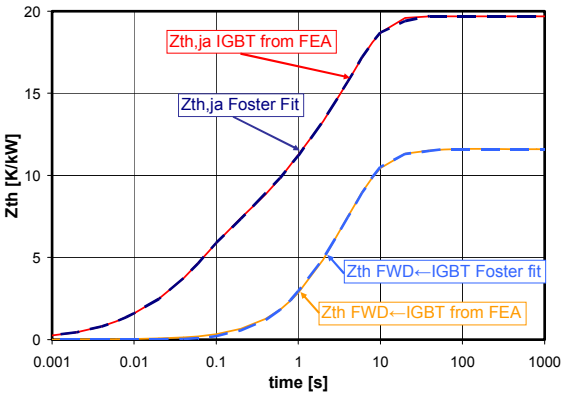

Fig. 11 Comparison between Z_{th} curves from FEA and their respective fits to Foster RC networks.

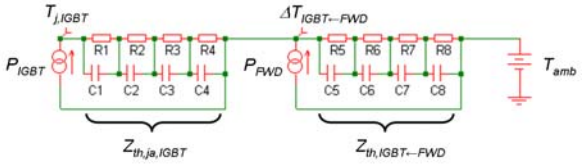

Fig. 12 Foster RC networks for determining T_j of the IGBTs including the temperature rise due to coupling to the FWDs.

Merging the RC-networks together one results in circuits as shown exemplarily for the IGBT in fig. 12. Both die types act as independent heat sources which are modelled as current sources.
The power losses in the dies for simulating the inverter operation shown below use sinusoidal half-waves to model the heat in the dies. This approach is based on the calculation methods used in IPOSIM [9]. It neglects the influence of the PWM ripple on the temperature. However, the temperature ripple would be only visible in the response of the direct heat path, not in the coupling because of the averaging effect due to the time delayed response in the convolution according to eq. 4.

4.2 Load cases with α=15000W/m²K

The losses used in the following simulations are taken from IPOSIM simulations. They are based on I_{rms}=1000A at output f_0=50Hz and I_{rms}=800A at f_0=1Hz. The switching frequency is held at f_s=1500Hz in both cases. The modulation is m=1. Finally, the ambient temperature is kept at 20°C.

4.2.1 f_0=50Hz, cosϕ=1

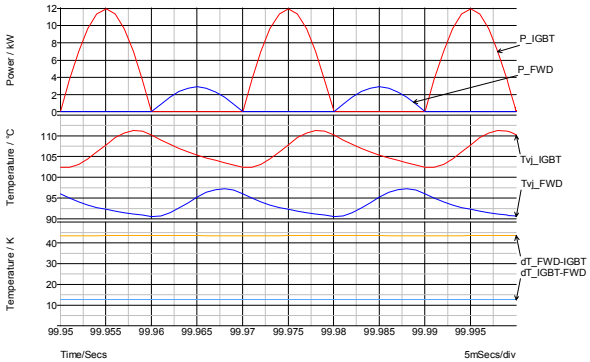

Fig. 13 Power profile and temperature of dies in HV-IGBT module during operation. The upper diagram shows the power loss profile. The middle contains the junction temperatures. The lower diagram shows the offset temperatures due to the mutual coupling.

4.2.2 f_0=50Hz, cosϕ=-1

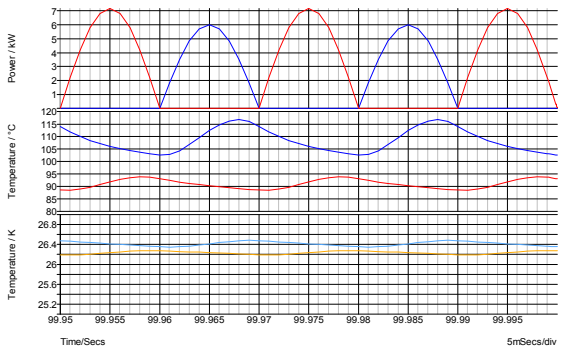

Fig. 14 Power profile and temperature of dies. Diagrams and colours are as in fig. 13.

4.2.3 f_0=1Hz, cosϕ=1

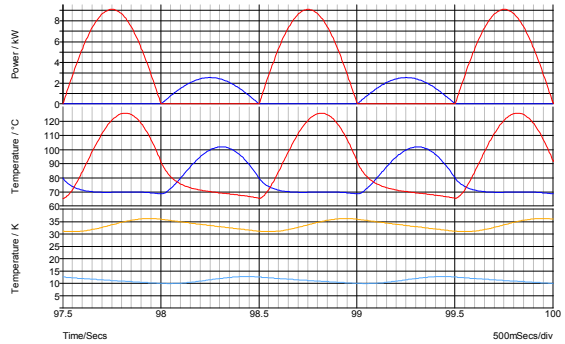

Fig. 15 Power profile and temperature of dies. Diagrams and colours are as in fig. 13.

At f_0=50Hz the temperature ripple of T_j is only governed by the direct thermal response. The coupling give only a homogeneous base with a ripple <0.15K. The situation at f_0=1Hz is different due to the visible influence of the crosstalk on T_j. The coupling ripple FWD←IGBT is about 5.5K, the absolute FWD temperature swing is 35K. The crosstalk thus flattens the ripple by roughly 15%. The IGBT ripple is affected by only 6%.

4.3 Load cases with α=1000W/m²K

Values of I_{rms}=450A at f_0=50Hz and I_{rms}=400A at f_0=1Hz were taken because of the larger $R_{th,ha}$. The switching frequency was held at f_s=1500Hz. The modulation was m=1, T_{amb}=20°C.

4.3.1 f_0=50Hz, cosϕ=1

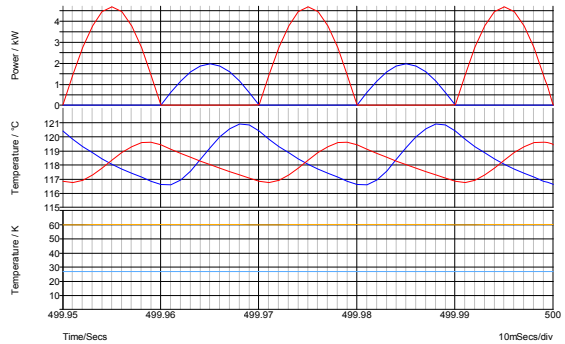

Fig. 16 Power profile and temperature of dies. Diagrams and colours are as in fig. 13.

4.3.2 f_0=50Hz, cosϕ=-1

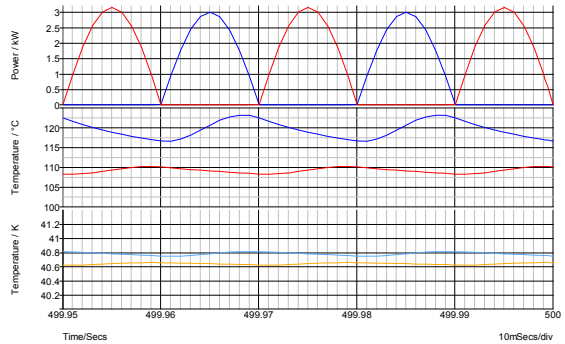

Fig. 17 Power profile and temperature of dies. Diagrams and colours are as in fig. 13.

4.3.3 f_0=1Hz, cosϕ=1

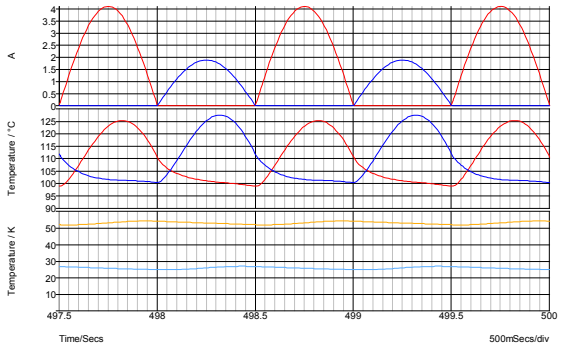

Fig. 18 Power profile and temperature of dies. Diagrams and colours are as in fig. 13.

The major difference between the water cooler from section 4.2 and the present case is the smaller absolute ripple in *Tj*. Additionally, the simulations suggest a slightly larger ripple for the FWD than for the IGBT in motor action cosϕ=1. Again, at f_0=1Hz, the crosstalk contribution on T_j flattens the ripple.

5 Summary

The paper deals with the description of the thermal crosstalk in HV-IGBT modules. First, the mutual coupling was studied on the die level. Due to the complexity a more global description was introduced. Within that the dependence of the crosstalk on the heat sink properties as well as the thermal grease layer was studied.
Compact RC-network models were deduced which allows for a study of the thermal coupling during inverter operation. At operation frequencies f_0=50Hz the crosstalk leads to a homogeneous temperature base whereas at lower frequencies (f_0=1Hz) a significant influence is found. This mainly flattens the ripple in T_j which is important e.g. during start of a vehicle. It may have impact on lifetime predictions.

The authors thank Peter Türkes for providing the Zth-fitting algorithms.

6 Literature

[1] M.J. Whitehead, C.M. Johnson, Proc. Electronics Systemintegration Technology Conference 2006, 1218
[2] U. Drofenik et al., Proc. PCIM Europe 2007
[3] T. Kojima, et al., R&D Review of Toyota CRDL, 39 (2004) 27; T. Kojima, et al., Proc. PESC 2006, 2048
[4] Y.C. Gerstenmaier, et al., IEEE Trans. Power Electr., 21 (2006) 45
[5] M. Ayadi, et al., Proc. Therminic 2005
[6] Ansys Inc., www.ansys.com
[7] F.W. Gutzwiller, T.P. Sylvan, Trans. AIEE, 79 (1961) 699
[8] Y.C. Gerstenmaier, G. Wachutka; Proc. EPE 1999
[9] IPOSIM, www.infineon.com