Thermal Modelling of Power Modules in a Hybrid Vehicle Application

Jonas Ottosson

Lund University

Doctoral Thesis in Electrical Engineering
Department of Measurement Technology and Industrial Electrical Engineering

2013
Abstract

Hybrid electric and full electric vehicles have attracted growing attention during the last decade. This is a consequence of several factors, such as growing environmental concerns, increasing oil prices and a strive for oil independency. Hybrid vehicles have proven to have significant potential to improve fuel economy and at the same time enhance the performance of the vehicle. For the hybrid cars to really penetrate the passenger car market, it is of vital importance that the cost of hybridization can be kept as low as possible and that the increased cost will be paid back within a reasonable time horizon. In addition to this, high reliability of the additional electrical drive system (EDS) needs to be ensured.

The main parts of the electrical drive system (EDS) are the three-phase inverter and the electric traction machine (ETM). The three-phase inverter, or commonly just referred to as the power electronics (PE), is often packaged into some kind of standalone housing. This is a simple and rather straight forward way of packaging, however it is not the preferred choice in a passenger car, where the number of places and the space to put the PE-assembly is often limited. Except for the high packaging volume, other drawbacks of using a standalone housing for the PE is that it requires expensive shielded high-voltage connectors and cables. An alternative solution is to integrate the power electronic inverter and the electric machine into one unit, sharing the same cooling system. This will reduce the packaging volume and at the same time omit the need for the expensive connectors and cables. The integration has other benefits such as, modularity and no separate housing for the inverter is required. All of the above mentioned benefits of the system integration lead to reduced costs of the system and in the long run a reduced price of the vehicle. However, integrating the PE and ETM into one unit makes the thermal design more difficult. This thesis focuses on determining the required cooling capacity of the cooling system, which is an important task, both from a reliability and cost perspective. Long lifetime or high reliability is important for customer acceptance.

Performing simulations of power inverter systems, where both the
electrical and thermal response is incorporated, often referred to as electro-thermal simulations, is a difficult task due to the different time scales of the two disciplines. If every switching instant is to be simulated, simulation times in the range of microseconds have to be used. Simulating inverter systems in a hybrid vehicle application, where driving cycles lasting for thousands of seconds normally is used, requires unreasonable simulation times. The thesis presents a method for solving this problem.

Several electro-thermal simulations are carried out in order to determine the cooling requirements and its effect on module reliability of the inverter. The foundation for these simulations is a simulation model of the electrical drive system, where the main focus is on the thermal model of the inverter itself. Finite element analysis (FEA) can provide highly accurate calculations at the expense of computational effort and simulation time. There is no easy way to incorporate FEA calculations into an electrical time domain simulation. The thesis presents a thermal model development procedure to derive simplified thermal models based on thermal impedances.

Thermal models for a range of power modules, including both Si and SiC based modules, are developed. Different types of assemblies with single-sided cooling are studied in this thesis. In addition to the different single-sided cooled module assemblies, one assembly with double-sided cooling is studied. The thesis shows that the type of layout and assembly greatly affects the thermal behaviour, and as a consequence the lifetime, of the power module.

In addition to different module assemblies and cooling options, the system is evaluated for different driving cycles and cooling medium temperatures, together with a comparison of using a fixed and variable switching frequency. The thesis shows that using a variable switching frequency has a significant impact on the cooling requirements.
Acknowledgements

First and foremost, I want to thank my supervisor Professor Mats Alaküla for the guidance during the work with this thesis.

I would also like to thank my former colleagues at BorgWarner (Haldex Traction) in Landskrona, special thanks to Daniel Hervén, Daniel Norlén, Gustaf Lagunoff and Pierre Pettersson who all have inspired me in different ways. The cooperation with BorgWarner has been very valuable to me.

I would like to thank the Department of Industrial Electrical Engineering and Automation at Lund University for all the interesting discussions. Special thanks to Getachew Darge and Bengt Simonsson for helping me out in the lab. I am also grateful to Dr. Per Karlsson for answering all my power electronics related questions. I would also like to thank my fellow PhD-students Jonas Johansson, Luyu Wang and Francisco Marquez, for good advice and collaboration. I am particularly indebted to Dan Hagstedt, my colleague and dear friend, for travelling with me on conferences and introducing me to the Hagstedts Fridays.

Many thanks to Fredrik Olne, one of my oldest and best friends, for having me as a lodger during my numerous trips to Skåne.

This work has been financed by BorgWarner and the Swedish Energy Agency, which is gratefully acknowledged.

Lastly, I would like to thank my family for all their love and support. A special thanks to my mum Carina for always being there for me, supporting me in all my pursuits. Lisen, one thing is true, this thesis would never have been finished without your support. Thanks for your patience, love and encouragement during these years. Eva-Li, you are my everything.

Göteborg, October 2013
Jonas Ottosson
Preface

When I write this, it is almost ten years ago since I started as a PhD student at the Department of Industrial Electrical Engineering and Automation, Lund University, in 2004. The project that I was working in back then was focused on energy management and control of electrical drives for hybrid electrical vehicles. The results of that project were published in a Licentiate thesis in 2007. In the same year I got an employment at Volvo AB and worked there for nearly two and a half year. In 2009 I got the opportunity to continue my PhD-studies. However, this was within a different project, focusing on the development of an electrically driven rear-axle for a hybrid car, where I came to focus on thermal modelling of power modules.
Contents

CHAPTER 1 INTRODUCTION ................................................................. 1
1.1 BACKGROUND ............................................................................. 1
1.2 MOTIVATION ................................................................................ 1
1.3 OBJECTIVES ............................................................................... 2
1.4 PREVIOUS WORK AND LITERATURE REVIEW ............................ 3
1.5 MAIN CONTRIBUTIONS ............................................................... 5
1.6 OUTLINE OF THE THESIS ......................................................... 6
1.7 PUBLICATIONS ............................................................................. 7

CHAPTER 2 POWER DEVICES ............................................................. 9
2.1 POWER DEVICE PACKAGING AND MATERIALS ......................... 9
2.2 POWER MODULE RELIABILITY .................................................. 17
2.3 HEAT TRANSFER .......................................................................... 27
2.4 POWER ELECTRONIC COOLING METHODS ................................. 28

CHAPTER 3 THERMAL MODELLING OF POWER MODULES ........ 31
3.1 THE HEAT CONDUCTION EQUATION ........................................... 31
3.2 SOLVING THE HEAT CONDUCTION ........................................... 32
3.3 TRANSIENT THERMAL IMPEDANCES ......................................... 36
3.4 MATHEMATICAL DERIVATION OF FOSTER BASED THERMAL
   MODELS ......................................................................................... 39

CHAPTER 4 SYSTEM MODELLING ...................................................... 45
4.1 SYSTEM MODEL ........................................................................... 45
4.2 POWER MODULES ........................................................................ 46
4.3 VEHICLE MODEL ........................................................................ 48
4.4 ELECTRICAL MACHINE MODEL ............................................... 49
4.5 POWER ELECTRONIC LOSS MODELLING ..................................... 52
4.6 THERMAL MODELS OF POWER MODULES ................................. 56

CHAPTER 5 SYSTEM SIMULATIONS .................................................. 63
5.1 SYSTEM SIMULATIONS AND EVALUATION CRITERIA ............... 63
5.2 DRIVING CYCLES ......................................................................... 63
5.3 CONVECTION COEFFICIENTS ...................................................... 64
Chapter 1

Introduction

1.1 Background

This thesis is a part of a project originally called eBAX (Swedish: elektrisk bakaxel), which overall aim was to develop an electrically driven rear axle for a passenger car. The name has changed over the years, and the most recent name is eTVD (Electrical Torque Vectoring Device. BorgWarner, former named Haldex Traction AB, had developed a first prototype. During that work, it was found out that more than 80% of the total system cost originates from the power electronics and the electric machine. Hence, the project was focused on three major parts, namely the power electronics, the electric machine and the overall optimization of the system.

1.2 Motivation

Hybrid electric and full electric vehicles attract growing attention. This is a consequence of several factors, such as growing environmental concerns, increasing oil prices and a strive for oil independency. According to some, we are rapidly approaching the top of the oil era, whereas some state that we have already passed peak oil. If this is true, the diminishing availability of crude oil will escalate the already high fossil fuel prices. Many countries, there among Sweden, have declared national actions plans to reduce the oil consumption significantly. The background of the desired reduction in oil use, differ between countries, but the two main concerns are the environmental effects and the strive for independency of oil import. Whatever the concern, economical, independency or environmental, how does hybrid vehicles fit in this context?

Hybrid vehicles have proven to have significant potential to improve fuel
economy without reducing the performance of the vehicle. Keeping or even enhancing, which often is the case, the drivability is an important issue in order to reach customer acceptance. Except for gains in the fuel economy, hybrid vehicles also demonstrate potential for reducing exhaust emissions. Most hybrid vehicles on the market today, are hybrid electric vehicles (HEV) with an internal combustion engine (ICE) and at least one electrical machine. Toyota was the first major automaker to introduce a commercial vehicle, the Toyota Prius. The commercialization in the recent years has mainly been possible due to advances in the battery and power electronics technologies. The HEVs mainly benefit from the following: down-sizing or “right”-sizing of the engine, still fulfilling the power demanded from the driver, recovering energy during deceleration by operating the electrical machine as a generator, eliminating the fuel consumption originating from engine idling by shutting the ICE off and operating the ICE more efficiently, since an extra degree of freedom is available to satisfy the power demand.

With this background given, it is clear why almost all car manufactures are developing HEVs today. However, for the hybrid cars to really penetrate the passenger car market, it is of vital importance that the cost of hybridization can be kept as low as possible and that the increased cost will be paid back within a reasonable time horizon.

1.3 Objectives

As was stated in the motivation for this thesis, there is a need from the car manufacturers to simplify and reduce the cost of hybrid drivelines, in order to make hybrid cars more attractive on the market. The hybrid vehicle, studied in this thesis, is assumed to be a plug-in hybrid electrical vehicle (PHEV) with an electrically driven rear-axle. The main parts of the electrical drive system (EDS) are the three-phase inverter and the electric traction machine (ETM). The three-phase inverter, or commonly just referred to as the power electronics (PE), is often packaged into some kind of standalone housing. This is a simple and rather straight forward way of packaging, however it is not the preferred choice in a passenger car, where the number of places and the space to put the PE-assembly is often limited. Except for the high packaging volume, other drawbacks of using a standalone housing for the PE is that requires expensive shielded high-voltage connectors and cables. An alternative solution is to integrate the
power electronic inverter and the electric machine into one unit, sharing the same cooling system. This will reduce the packaging volume and at the same time omit the need for the expensive connectors and cables. The integration has other benefits such as, modularity and no separate housing for the inverter is required. All of the above mentioned benefits of the system integration lead to reduced costs of the system and in the long run a reduced price of the vehicle. However, integrating the PE and EM into one unit makes the thermal design more difficult. An inverter able to deliver a peak power output of 90kW, with an average efficiency of 97%, produces approximately 2.7 kW of losses, which needs to be handled by the cooling system. Determining the required cooling capacity of the cooling system is hence an important task, both from a reliability and cost perspective. Long lifetime or high reliability is important for customer acceptance. A detailed analysis of the cooling requirements for the PE, prevents the cooling system and the power modules from being overrated. Hence, based on what has been discussed above the main objectives of this thesis are:

- Investigate if it is possible to integrate the power electronics into the rear drive unit, by deriving guidelines for the cooling system in terms of required convection coefficients and maximum cooling liquid temperatures.
- Determine the actual cooling requirements for the power electronics, looking at different driving cycles and module assemblies.
- Develop thermal models that can be used in system simulations.
- Study how the requirements are changed if high temperature devices like SiC BJT:s are used.
- Study how the reliability is affected by the cooling requirements.

1.4 Previous Work and Literature Review

This section is devoted to a literature review of previous work within the area of thermal modelling of power modules. The literature is divided into three different categories. Some of the papers fit in more than one category, however the actual categorization is based on what the author finds most interesting about the specific paper.

Thermal Modelling of Power Modules

Power modules used in hybrid vehicle applications have to be able to
handle high currents. This is solved by parallel connection of multiple IGBT and diode chips within the same power module. This complicates the thermal modelling since thermal coupling within the module has to be considered [40] [54] [55]. There are many ways of building a thermal model of a power module. The finite element method can provide highly accurate calculations at the expense of computational effort and simulation time [53]. There is no easy way to incorporate FEM calculations into an electrical time domain simulation. Simplified thermal models are therefore often used and the most frequently used models are based on thermal networks. There are two different main types of thermal networks, the Cauer and the Foster network. The former is a simplified physical model describing the thermal behaviour of the power module, whereas the latter is a behavioural model usually derived from some kind of curve fitting procedure. The Cauer model can describe three dimensional heat flow [33], however it requires a rather fine discretization of the power module in order to capture the thermal coupling between its power devices. In order to derive a Foster model, results from either FEA simulations or measurements need to be available. The reason is that it is a behavioural model and therefore needs to be based on known behaviours. Normally some kind of curve fitting procedure is used to extract the parameters of the model. Several Foster models can be arranged into a thermal impedance matrix describing the thermal coupling between the devices within the power module [6][35][38][40]. Thermal impedance matrices are used extensively throughout this thesis, due to its ease of extraction, implementation and model accuracy. Another type of thermal model described in [29]-[32] is based on solving the heat conduction equation with a Fourier series solution method. The method is according to [32] faster than the FDM with an acceptable level of accuracy and can easily be programmed in the Matlab/Simulink environment.

**Electro-thermal System Simulations**

Performing simulations of power inverter systems, where both the electrical and thermal response is incorporated, is a difficult task due to the different time scales of the two disciplines. If every switching instant is to be simulated, simulation times in the range of microseconds have to be used. The thermal time scale is higher and normally in the millisecond range, depending on the size of switching devices. Simulating inverter systems in a hybrid vehicle application, where driving cycles lasting for thousands of seconds normally is used, would require unreasonable
1.5. Main contributions

simulation times. In order to speed up the simulations, decrease the simulation time and reduce the required memory space, semiconductor power losses are pre-computed and stored in look-up tables [42]. In [42] the power losses are computed from datasheets and a rather simple model of the EM is used. The authors in [16] and [43] [44] uses compact device models to generate the power loss look-up tables. The combination of using look-up tables for the power losses and simplified thermal models make it possible to evaluate many different system set-ups without simulation times becoming excessive. In [6] different system set-ups in terms of battery voltage, ambient temperatures and cooling options (convection coefficients) are studied.

Reliability of Power Modules
The method used in this thesis to obtain a reliability estimation of the power module lifetime consists of three different steps. The first step is to carry out electro-thermal simulations in order to obtain a thermal evolution of the chip temperatures. The second step is to decompose the thermal evolution into frequencies of different combinations of temperature swing and the associated mean temperature at which the temperature swing occurs. This is done by applying a cycle counting algorithm to the junction temperature evolution. The most commonly used algorithm is the rainflow algorithm [22]-[23], often used within the area of material fatigue analysis, however other algorithms can also be used [21] [47]. As the extraction of the thermal cycling is obtained, the result is used together with some kind of reliability model. The LESIT model is a reliability model that was presented almost fifteen years ago [15]. However, it is still used as starting point of reliability estimation and often used as comparison for new improved packaging technologies [16] [17] [18]. A more recent reliability estimation model is the CIPS 2008 model, which was presented in [19]. The model is derived from extensive testing of different power modules. This model has more parameters than the LESIT model and these additional parameters are related to the power module type and layout. In [20] a lifetime model based on plastic strain is derived and further developed in [21], including creep behaviour of solder layers.

1.5 Main contributions

The results and contributions of this thesis are presented in Chapter 6 and discussed in Chapter 7. The main contributions of the thesis can
summarized as follows:

- The effectiveness of using thermal models based on thermal impedances is proven.
- The importance of taking thermal coupling between the devices within a power module into account is pointed out and exemplified.
- Different module assemblies are compared in terms of cooling capability.
- Single- and double-sided cooling are compared both on a module and complete system level.
- The cooling requirements for 300 and 400A modules are determined for a wide range of commonly used driving cycles.
- It is shown that the cooling requirements can be significantly reduced by using a variable switching frequency.
- Using SiC based power modules reduces the requirements on the cooling significantly.
- An analysis of what impact a hill-hold feature has on the cooling requirements is studied.
- A reliability analysis on system level is presented.

1.6 Outline of the Thesis

This chapter is devoted to a brief outline of the thesis and a short introduction to each chapter is given. The thesis is focusing on thermal modelling of power modules and an introduction to packaging of power devices, with focus on power modules, is given in Chapter 2. How to model power module reliability and a short introduction to the discipline of Heat Transfer is also included in Chapter 2.

There are several ways of modelling a power module thermally, and a brief description of the most common ways of solving the heat conduction equation is given in Chapter 3. The main focus is however on the concept
of thermal impedances and a mathematical derivation of thermal impedance based thermal model is presented.

The thermal simulations presented in this thesis are based on a simulation model of a hybrid system for a passenger car. Each of the components of the full system model is described in detail in Chapter 4.

The system simulations and the criteria for evaluation of the results are described in Chapter 5. Full system simulations are carried out for a range of different conditions, such as different driving cycles, cooling requirements, module assemblies, cooling medium temperatures and switching frequency strategies, in order to obtain a comprehensive view of the system.

The results of the system simulations are in Chapter 6. The results are divided into two separate sections, where the first one is devoted to determining the actual cooling requirements of different power modules and the second one presents a reliability analysis of the modules.

The thesis ends with some concluding remarks and suggestions for future research in Chapter 7.

1.7 Publications

Parts of the work presented in this thesis have been presented in the following publications:


Chapter 2

Power devices

2.1 Power Device Packaging and Materials

Package Types
There are many different package types for power switches. The basic features of the package depend on the actual definition of what a power device package is. One possibility is to define the package as everything around the actual die. Two obvious basic features are then to support and protect the rather fragile die and serve as a supporting structure for the electrical connections. However, a good power device package should also provide good thermal performance in order to keep junction temperatures as low as possible. An easy and cost effective production is also important. More specific features are to ensure a long lifetime of the device and to minimize undesirable electrical properties, such as parasitic resistance, capacity and inductance [1].

The most demanding challenges for power device packages, from a reliability perspective, are to withstand the induced stress in interconnection of different materials due to thermal cycling. Every material within the device has a coefficient of thermal expansion (CTE), which is a measure of the degree of expansion of the material for a change in temperature. As a device is heated up, either externally by the surroundings or internally by its own losses, the different materials will expand differently causing bending within the structure. The stress induced by the difference in CTE will over time destroy the device.

Discrete Packages
There are mainly three different package types used for power switches
and the type is tightly coupled to the power level of the device. In low and high power applications discrete components are normally used. However, their packaging assemblies are totally different, which can be seen in Figure 2.1.

The most common type of package for the low power switches is the TO-package. This type of package basically consists of a copper baseplate on which the silicon chip is soldered, see Figure 2.1(a). Bond wires connect the silicon chip to the lead frames. A mold is added to support the bond wires and protect the silicon chip from the surrounding environment. Other packaging designs have evolved from the TO-package, where development in order to increase reliability and to minimize electrical parasitics has been done.

When it comes to really high power applications, discrete components are also used and the dominating packaging type is the capsule. An example of the capsule packaging type and a schematic overview of the internal construction can be seen in Figure 2.1. The reasons for using capsules in the high power range are its compact design, the possibility to cool the device on both sides, the lack of bond wires and the limited number of
interfaces between materials with different thermal expansions [1].

**Standard Power Module Assemblies**

A power electronic module, or often shorter power module, is a packaged device that can contain different setups of switches and diodes. Common setups are half-bridges and full three phase inverters. Several switches and diodes are often parallel connected in order to increase the current capabilities of the module. Both MOSFET:s and IGBT:s are used in power modules, and the choice of which one to use depends on the application. There are power modules, often referred to as intelligent power modules (IPM), where peripheral electronics such as the driver and protective circuits are built in to the same package as the power devices. Common protection features are over current and short circuit protections, together with temperature monitoring.

A cross sectional view of the most common power module packaging structure is presented in Figure 2.2(b). The figure also shows an example of an IGBT power module from Infineon [2], containing a full three phase inverter.

![Figure 2.2](image)

Figure 2.2  Power module from Infineon [2] (a) and a cross sectional view (b) of the layered power module structure.

The silicon chips are soldered onto an isolation layer that consists of a direct-bond-copper (DBC) structure, centred round a ceramic substrate. The ceramic isolation layer isolates, electrically, the current carrying parts from the rest of the module structure. Bond wires, normally made out of aluminium, are used for interconnections on the DBC. The bond-wires are also used to connect the current carrying part to the power module terminals. The DBC is soldered onto the module baseplate, which both
acts as a mechanical support, for the rather fragile DBC, and as a heat spreader, increasing the heat dissipating area. The module itself is often mounted on a heatsink that dissipates the heat generated within the module. Both the baseplate and the heatsink have surfaces that, in a microscopic scale, are not completely smooth. These irregular surfaces prevent good thermal contact and a large part of the contact surface is separated by a layer of interstitial air [3]. Since air has low thermal conductivity, a Thermal Interface Material (TIM) is used to enhance the contact. There exists a large variety of TIMs and the most common ones are commonly referred to as thermal grease. The thermal grease is applied to one of the surfaces and fills the micro voids with grease instead of air, greatly reducing the thermal resistance of the contact layer. However, it should be pointed out that the TIM layer is prone to degrade over time and can drift away as a consequence of the bending of the baseplate, caused by the thermal cycling.

There are two variants of the classical power module structure presented in Figure 2.2 that are getting more common. One is to omit the baseplate and mount the substrate directly on the heatsink, while the other is to use the baseplate itself as a heatsink. Omitting the baseplate has the benefit of reducing the stress in the solder joint between the silicon chips and the substrate [4]. The reason is that the substrate is mounted closer to the heatsink and hence will have a lower temperature, resulting in less stress between the chip and substrate. In addition to this, the large solder joint between the substrate and the baseplate is eliminated. This joint is a weak point and is often prone to failure [5]. A drawback is however that the thermal inertia is lowered and this can cause higher junction temperature swings in applications where frequent low frequency high power pulses occur [4]. The second variant mentioned is to omit the heatsink and let the baseplate itself act as a heatsink, often referred to as direct cooling. The heat spreading effect of the heatsink is then lost, but the effective baseplate area can be increased by using a pin fin structure [6]. A benefit of using direct cooling is that the thermal grease can be removed.

**Power Module Assemblies for Enhanced Cooling**

In addition to the more common power module assemblies mentioned in the previous section, more innovative solutions exists where the aim is to enhance the cooling of the device. Standard power module assemblies are normally cooled on the non-wire bonded side of the assembly. Omitting
the bond wires, by using other interconnection solutions, makes it possible to cool the semiconductor devices from both sides. Some examples of different interconnection solutions where bond wires are not used are low temperature joining, flip chip solder balls, press pack, DirectFET, power overlay and embedded power technology [49]. The different bond wire free interconnection solutions are used together with different types of double-sided cooling techniques. The design solution presented in [50] – [52] is based on using two DBC:s, where the additional substrate is placed on top of the device, see Figure 2.3.

![Figure 2.3 Power module based on two DBC:s with double sided cooling.](image)

Using two substrates makes it possible to extract heat from both sides of the device, which, compared to a version with single sided cooling, has higher current carrying capability due to a reduced thermal resistance. From Figure 2.3 it can be seen that the proposed solution does not have any baseplate, which from a reliability perspective is beneficial since the solder layer between the substrate and baseplate is prone to degrade over time. The solution with two substrates and no baseplate is studied more in detail in later chapters of the thesis. Thermal simulations of the design solution are carried out and compared to a more standard solution with single-sided baseplate cooling.

There are several advantages of using double-sided cooling, where one have already been mentioned, a decreased thermal resistance, since heat is extracted from two sides and decreased number of layers within the module structure. The baseplate is the heaviest part of the power module assembly and omitting it has a significant impact on the total module weight. Another advantage is that the structure carrying the cooling liquid
can be made from any material, since it is not actually conducting any heat [50]. In the proposed solution the assembly is clamped in between the structures carrying the cooling liquid. The clamping force that needs to be applied is quite low since it only needs to be high enough to create a proof sealing. Due to the low clamping force, no mechanical stress is transferred between module assembly and the cooling structure [50].

There are mainly three drawbacks of the module assembly presented in Figure 2.3. The two substrates are in close vicinity to each other, around 100μm, and have a high electric potential difference which causes a high electrical field. This can be solved by etching of the DBC in exposed areas [52] or by using spacers to increase the clearance. Another drawback is that it is difficult to coincidentally bond the silicon chips to the bottom and top substrates [52]. Even though omitting the baseplate has its advantages, it can actually have a negative impact on the cooling performance, since its heat spreading effect and its thermal inertia is lost. Advantageous or not, is a question of the cooling technique applied to the module assembly. This is further investigated in Chapter 5.5.

**Power Module Packaging Materials**

Power modules consist of several different materials and the most common ones are listed in Table 2.1. Included in the table are also the key parameters from a power module packaging and reliability perspective.

The most common material for the chip of commercially available power devices is Silicon. This has been the dominating material since the introduction of semiconductor devices in the middle of the twentieth century. The main reasons for the dominating use of Silicon are its simple processing, useful temperature range and low price. In later years, devices based on Silicon Carbide (SiC) have become more and more common. The reason for using SiC is its wider band gap, approximately 3 times higher than for Silicon, which results in much smaller intrinsic carrier concentrations [7]. Table 2.1 shows that the thermal conductivity for SiC is two to three times higher than that for silicon. This together with the lower intrinsic carrier concentration makes the SiC material more suitable for high temperature applications.
For attaching the chips to the substrate and the substrate to the baseplate, some kind of bonding material has to be used. The most common category of bonding materials for power semiconductors are solders, which are alloys of two or more materials. The most common solders today are variations of Tin/Silver (Sn/Ag) alloys. These solders have replaced the Tin/Lead (Sn/Pb) based solders used in the past, in order to meet new environmental requirements. The lead free solders have higher melting points and correspondingly higher reflow temperatures. Hence, using lead-free solders makes the soldering process more critical. The drawback of using traditional solders and the soft soldering technique, is that it cannot be used for junction temperatures above 125°C. The reason for this is that the solder joint is getting more sensitive to strain as the melting point of the solder is approached. Higher chip temperatures, approaching 175°C, will dramatically reduce the reliability of devices due to the soft soldering technique. When comparing different bonding materials a term called homologous temperature is often used and it is defined as the ratio between the maximum operating temperature and the melting temperature (in the Kelvin temperature scale) [8]. High ratios will result in poor mechanical strength and creep under stress. Ratios lower than 40% have very little impact on material properties. In the range between 40 to 60%,
Chapter 2. Power devices.

material properties start to change and the solder is getting more sensitive to strain and for ratios above 60% materials are considered unable to bear engineering loads in a structure. [4]. The homologous temperature for the most commonly used soft solders today, assuming two different operating temperatures, 125°C and 175°C, are approximately 80 and 90%, respectively. These high ratios lead to fatigue cracks in the solder layers when the device is subjected to thermo-mechanical loading. This has resulted in the development of new soldering techniques, offering stronger chip die attach with a higher creep resistance. The two most promising techniques are transient liquid phase soldering (TLPS) [8]-[9] and silver sintering [9]-[11]. The homologous temperatures for these techniques are 52-65% [8] and 35% [4], respectively.

The reason for using a substrate is mainly to isolate the chip electrically from the baseplate. The substrate also serves as a supporting structure for the power module circuitry. Using a substrate is also beneficial for the reliability since the chip and substrate has better matching of CTEs, than chip and baseplate, which can be seen in Table 2.1. A substrate should offer high thermal conductivity, matching CTE with the chips, high dielectric strength and a low cost [12]. There are basically three substrate material candidates, Alumina (Al₂SO₃), Aluminium Nitride (AlN) and Silicon Nitride (Si₃N₄), all having its pros and cons, see Table 2.1. The most common one in power semiconductor applications is the Al₂SO₃. It offers low cost and average mechanical, thermal and electrical characteristics. The drawbacks are its relatively low thermal conductivity and rather high thermal expansion mismatch with Si [12]. A better option could be AlN which has a much higher thermal conductivity, a better thermal expansion matching with silicon and breakdown voltage capabilities. However, ceramic insulators based on AlN are considerably more expensive, a factor of two to three, than those based on Al₂SO₃. The last candidate is the Si₃N₄, which has good thermal conductivity and a very good thermal expansion match with silicon. It also has almost twice as high mechanical fractural toughness compared to Al₂SO₃ and AlN. This makes it suitable for applications where it is beneficial to omit the supporting baseplate. A drawback is the higher price and limited number of suppliers [12].

The reason for using a baseplate in the power module package is to offer mechanical support for the, sometimes rather fragile, substrate. The
baseplate acts as a thermal inertia, absorbing heat during power transients within the chip. The baseplate conducts the heat to the heatsink, and has a heat spreading effect which increases the effective heat dissipating area. Copper is the most commonly used material for baseplates due to its high thermal conductivity and relatively low cost. The drawback is its thermal expansion mismatch with the substrates. For high performance applications, where AlN substrates are often used, the preferred baseplate material is the Aluminium/Silicon carbide (AlSiC) metal matrix composite [1]. The reason for using AlSiC in these types of applications is the better matching in thermal expansion between the substrate and the baseplate.

2.2 Power Module Reliability

Failure Mechanisms
Power modules for traction in hybrid electric vehicles are subjected to several thermal cycles due to self-heating, where the losses generated within the device heats the module itself. The power module consists of several layers of several different materials, for which the coefficient of thermal expansion (CTE) can vary significantly. This variation in CTE induces stresses in the materials and in the interconnection of the different materials, as they are subjected to thermal cycling. The fatigue stress is a consequence of the bi-metallic effects and the actual level of stress depends on the thermal excursion or temperature swing. The most common failure mechanisms in power modules are bond wire fatigue, bond wire heel cracking, aluminium reconstruction and solder fatigue and solder voids [13]-[14]. The most dominant of these failure mechanisms is bond wire lift-off. This failure mechanism is mainly caused by the mismatch of the CTE for the aluminium wire and silicon chip. Repetitive thermal cycles induces shear stress on the interface between the bond pad and the wire and bending or flexure of the bond wire itself. Normally, several bond wires are connected in parallel to increase current capability. If a bond wire breaks, this results in an increased resistance and a redistribution of the current within the device. This will accelerate the failure rate of the remaining bond wires. The lift-off failure mechanism is also accelerated by solder fatigue. Solder fatigue increases the thermal resistance between the chip and the convective surface, and consequently causes the junction temperature to increase.
Reliability Estimation Models

The reliability for the power modules studied in this thesis is estimated using two different reliability estimation methods. Both of the methods are developed through testing of a large number of devices, where the results serve as a base for fitting a function estimating the number of cycles to failure. Hence, the estimation models are based on statistical analysis of the device testing results. The first model studied is often referred to as the LESIT model. Several different modules from different suppliers were tested for a wide range of thermal conditions, consisting of different junction temperature swings at different mean junction temperatures. The results from these tests were first presented in [15]. Even though this model is relatively old, it has still in recent years been used as a starting point for power device reliability estimation [16]-[18]. The number of cycles to failure, \( N_f \), can be computed as

\[
N_f = A \cdot \Delta T_j^\alpha \cdot e^{\frac{E_a}{k_B T_m}}
\]  

(2.1)

where \( k_B \) is Boltzmann’s constant, \( E_a \) is the activation energy and \( A \) and \( \alpha \) are fitting parameters. The temperatures \( \Delta T \) and \( T_m \), which are the temperature swing and the absolute mean junction temperature, respectively, are the dependent variables that determines the number of cycles until failure. The LESIT model is actually a combination of the Coffin-Manson law and an Arrhenius factor. The Coffin-Manson law states that the number of cycles until failure is a function of the plastic strain, which in turn is a function of the temperature swing [13]. In [15] it was shown that the number of cycles until failure, as a function of the temperature swing, appeared as straight lines when plotting them in a double logarithmic scale. As the mean temperature changed, the lines were shifted but still parallel, which indicates an exponential dependency. Hence, an Arrhenius factor was added to account for this thermally activated shift. It should be pointed out that the LESIT model was derived for mean temperatures ranging from 60 to 100°C, with temperature swings between 30 and 80°C. Hence, some care must be taken when the model is used for temperature combinations outside this range. Using the LESIT model for more recent power modules has proven to give a somewhat conservative estimation of the number of cycles to failure [17].
The second model used in this thesis was presented in [19], and is often referred to as the CIPS 08 model [1]. As for the LESIT model, this model is derived from extensive testing of different power modules. However, this model has more parameters and these are related to the power module type and layout. The number of cycles until failure, \( N_f \), can be computed as

\[
N_f = K \cdot \Delta T_j^{\beta_6} \cdot e^{\frac{\beta_1}{t_{on}}} \cdot I^\beta_4 \cdot V^\beta_5 \cdot D^\beta_6
\]  

where \( K \) and \( \beta_1 - \beta_6 \) are fitting parameters chosen according to [19]. The model contains other parameters, such as the heat-up time, \( t_{on} \), the current per bond stitch, \( I \), the voltage range of the device, \( V \), and the bond wire diameter, \( D \). Both the LESIT and the CIPS08 model have the temperature excursion, \( \Delta T_j \), as a dependent variable. The CIPS08, however, uses the starting point of the temperature swing, \( T_{low} \), as a parameter instead of the mean temperature. Apart from this it requires an estimate of the heat-up time associated with the temperature swing, see Figure 2.4 for a definition of the variables. The reason for including the heat-up time in the model is to capture the fact that the different layers of the power module have different time constants. The top layers, and the associated failure mechanism of bond wire lift-off, have a time constant in range of seconds, whereas the solder layer between the substrate and baseplate has usually a time constant at least ten times higher. Hence longer pulses will not only contribute to the bond wire lift-off, but also solder degradation.

![Figure 2.4](image)

Figure 2.4 Definition of the heat up time and the different temperatures used in the LESIT and CIPS08 models.
Chapter 2. Power devices

It should be pointed out that the CIPS08 model was derived for a range of each parameter in Equation 2.2 [19]. Hence, some care must be taken when interpreting results derived from the model with parameters outside of the validated range.

Both models presented above are used in the estimation of the lifetime for the modules studied in this thesis. It could therefore be of interest to compare the two models. Figure 2.5(a) shows a comparison of the number of cycles until failure for the LESIT and CIPS08 model. The comparison is presented as a ratio, \( r_{N_f} \), of the number of cycles until failure for different temperature swings and mean temperatures, computed according to

\[
r_{N_f} = \frac{K \cdot \Delta T_j^\beta_e \cdot e^\frac{\beta_e}{E_o} \cdot I_j^\beta_e \cdot V_j^\beta_e \cdot D_j^\beta_e}{A \cdot \Delta T_j^\alpha \cdot e^\frac{E_o}{\Delta T_j}}
\]

The on time for the CIPS08 model is, in Figure 2.5(a) assumed to be 1s. Figure 2.5(b) shows for what combinations of the temperature swings and mean temperatures the LESIT model estimates a higher number of cycles until failure than the CIPS08 model. This comparison is presented for three different on times, 1, 5 and 15s, respectively.

Two important things can be seen in Figure 2.5(a). The first one is that the CIPS08 model estimates a higher lifetime for almost all combinations of the temperature swings and mean temperatures. For this particular case, where ton = 1s, the CIPS08 model estimates, at the most, a 20 times higher lifetime. The ratio is decreased as ton is increased and is at the most 5 for ton = 15s. Figure 2.5(b) shows in what range the ratio computed according to Equation 2.3, is less than one, i.e. where the LESIT model is estimating a higher lifetime. The ratio is computed for three different on times, 1, 5 and 15s. The result shows that the LESIT model estimates higher lifetimes for a range of combinations where the temperature swings and mean temperatures are low and this range is growing as the on-time is increased. Hence, for mission profiles consisting of shallow cycles at low mean temperatures, the LESIT model is estimating a higher number of missions until failure.
Linear Accumulated Damage

A temperature profile consists of several temperature swings at many different mean temperatures. The reliability models presented here gives the number of cycles to failure for a specific operating point. Hence, a relation relating the different temperature swings and their associated number of cycles to failure must be used. The relation that is often used in literature is the Miner’s rule [16], [20]-[21]. This relation can, for two variables, be written as [16]

\[
C = \sum_{i,j} \frac{n_{i,j}}{(N_f)_{i,j}}
\]  

(2.4)

where \(C\) is the accumulated damage, \(n_{i,j}\) and \((N_f)_{i,j}\) is the number of cycles and the number of cycles until failure for the \(i\)th mean temperature and the \(j\)th temperature excursion. When the accumulated damage, \(C\), is close to one failure occurs. For any mission profile there will be different combinations of temperature swing and mean temperatures. Miner’s rule can be seen as the linear combination of the induced stress portions for the
different thermal states. Equation 2.4 can be directly used in the case of reliability estimations with the LESIT model. However, for the CIPS 08 model, it needs to be extended. Miner’s rule can easily be extended with more variables, since it is independent of the exact nature of the lifetime model [1]. The CIPS 08 model requires one more variable, the heat up time, \( t_{\text{on}} \). When using Miner’s rule, it is assumed that the accumulation of damage is linear. Hence, no respect is given to the order of the thermal cycling and the associated induced stress. However, as is pointed out in [16], the relation may still be used to derive a single value representing the accumulated damage of the device.

**Thermal Cycling**

The output of the system simulations in the coming chapters are, among other things, the junction temperatures for the IGBT:s and diodes in the power module. The thermal cycling, imposed by the operation of the electrical machine, serves as an input for the reliability study. Figure 2.7 shows the IGBT junction temperature evolution for a well-known driving cycle named US06, for two different convection coefficients, 1 and 10 kW/(m\(^2\)°C).

![Figure 2.6 IGBT junction temperature evolution for the US06 cycle for two different convection coefficients, 1.0 to 10 kW/(m\(^2\)°C).](image)
Figure 2.7 Vehicle speed and the driver requested wheel torque. IGBT junction temperature rise as a function of EM torque and speed for one electrical revolution (c).

The junction temperature excursion depends on the load conditions of the electric machine. The IGBT temperature rise of the IGBT chip, computed for one electrical revolution, is presented in Figure 2.7. The figure shows that the most demanding operating points, in terms of temperature increase, are the ones with high torque and low speeds. High torques in combination with low speeds implies high currents and relatively low modulation frequencies compared to thermal time constant of the power module heatsink, which results in a high temperature swing. The figure also shows that operating the machine close to the maximum torque in the field weakening region results in high temperature excursions. The reason is that field weakening operation requires both high current and high voltage levels, resulting in high losses for the switches.

**Cycle Counting**

The device experiences an irregular, or almost random, thermal cycling with different combinations of thermal excursions and mean temperatures. Cycle counting methods can be used to decompose these irregular thermal
evolutions into frequencies of these dependent variable combinations. Hence, cycle counting is used to determine the number of cycles of each combination of the dependent variables in the Miner’s rule. The variables that need to be considered for the LESIT model are the mean temperature and the temperature swing or thermal excursion. The CIPS08 model also requires the temperature excursion, but uses the starting point of the excursion instead of the mean temperature. Apart from this it requires an estimate of the heat-up time associated with the temperature excursion.

**Rainflow algorithm**

One of the algorithms used to extract the mean temperature and the temperature swing from the time evolutions of the junction temperature obtained from the full system simulations, is the rainflow algorithm. This algorithm is widely used for estimating stress/strain hysteresis loops within the area of fatigue life estimation [22]-[23]. Figure 2.8 shows an example of a thermal cycle and how the rainflow algorithm decomposes the cycle into temperature swings. Compare with Figure 2.9 for other cycle counting algorithms. The result of the rainflow algorithm is given in Table 2.2.

![Figure 2.8](image)

*Figure 2.8* An example of how a cycle is decomposed into its subcycles by the rainflow algorithm.
2.2. Power Module Reliability

The output of the rainflow counting algorithm is a matrix containing the number of occurrences of each combination of thermal excursion and temperature mean, together with the cycle period. The main difference of the rainflow algorithm compared to other cycle counting algorithms is that the rainflow algorithm really extracts the deep cycles, see Table 2.2 for a comparison.

Other cycle counting algorithms

There are many ways of extracting the thermal cycles of a junction temperature profile. Six different cycle counting methods are described below [21][47].

- “Rising edge”: The height of the rising edge within a cycle is used as the amplitude for the whole cycle (see Figure 2.9(a)).
- “Falling edge”: The height of the falling edge within a cycle is used as the amplitude for the whole cycle (see Figure 2.9(b)).
- “Mean edge 1”: Computed by taking the mean of the Rising edge and Falling edge (see Figure 2.9(c)).
- “Maximum edge”: The maximum height of the rising and falling edge within a cycle is used as the amplitude for the whole cycle.
- “Minimum edge”: The minimum height of the rising and falling edge within a cycle is used as the amplitude for the whole cycle.
- “Mean edge 2”: Computed by taking the mean of the Maximum edge and Minimum edge.

The cycle counting method called Mean edge 1 and Mean edge 2 gives the same result. The main difference between the cycle counting methods presented above and the rainflow algorithm is, as has been mentioned before, that the later finds the deeper cycles within a temperature profile. This can be seen in Figure 2.9 and the results presented in Table 2.2.
Figure 2.9 Three examples of cycle counting methods, "Rising edge" (a), "Falling edge" (b) and "Mean edge I" (c).

Table 2.2 Comparison of cycle counting algorithms.

<table>
<thead>
<tr>
<th></th>
<th>Cycle 1</th>
<th>Cycle 2</th>
<th>Cycle 3</th>
<th>Cycle 4</th>
<th>Cycle 5</th>
<th>Cycle 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainflow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ</td>
<td>4</td>
<td>2</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Mean</td>
<td>-1</td>
<td>2</td>
<td>1</td>
<td>0,5</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>N cycles</td>
<td>1</td>
<td>1</td>
<td>0,5</td>
<td>0,5</td>
<td>0,5</td>
<td>0,5</td>
</tr>
<tr>
<td>Mean cycle</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ</td>
<td>4</td>
<td>6</td>
<td>4,5</td>
<td>7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>-1</td>
<td>2</td>
<td>0,75</td>
<td>0,5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>N cycles</td>
<td>1</td>
<td>1</td>
<td>1,0</td>
<td>1,0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The cycle counting methods presented here is sensitive to temperature profiles containing ripple in the temperature, since these superimposed shallow cycles are extracted, resulting in the deeper cycles being overlooked [47]. This is a major drawback, since it is the deeper cycles that have the largest impact on the power module lifetime. One way to
partly overcome the sensitivity of the presented cycle counting algorithms, is to first filter the extracted temperature profiles. The rainflow algorithm is, due to its robustness, exclusively used for cycle extraction in the thesis.

2.3 Heat Transfer

The main purpose of any kind of cooling application for power electronics is to remove heat generated in the device chips to prevent the junction temperatures from reaching destructive levels. Heat can be defined as “the form of energy that can be transferred from one system to another as a result of a temperature difference” as stated in [24]. There are three basic mechanisms or modes for heat transfer, conduction, convection and radiation, and they are briefly described in the following section together with an overview of different power electronic cooling options.

Modes of Heat Transfer

*Conduction* is a heat transfer mechanism where energy is transferred in solids and stationary gases or liquids. The energy transport in solids is an effect of molecule vibrations and energy transport of free electrons, whereas the transport in liquid and gases is an effect of collision and diffusion of the actual molecules [24]. The heat transfer rate, $\dot{Q}$, through a plane layer due to conduction can be expressed as

$$
\dot{Q} = -\lambda A \frac{dT}{dx} \quad [W]
$$

(2.5)

where the constant $\lambda$ is the thermal conductivity, which is a material property describing the materials ability to conduct heat, $A$ is the heat transfer area and $x$ is the direction of the heat conduction. The relation above is called the one dimensional form of Fourier’s law of heat conduction. The rate of heat flux is proportional to the thermal conductivity and the temperature gradient across the layer. The minus sign is motivated by the fact that heat is always transferred in the direction of decreasing temperature. The dominating mode of heat transfer within power semiconductor devices is conduction. The main part of the losses generated in the device chip is transferred through the different layers down to the heatsink, where the heat transfer mechanism of convection finally removes the excess heat.
**Chapter 2. Power devices.**

*Convection* can be defined as heat transfer to a fluid or gas, which has a bulk motion [24]. The heat transfer is a combination of conduction and energy transport due to the movement of the fluid or gas. Convection is normally divided into two main categories, natural and forced convection. Fluid motion due to buoyancy forces, induced by the differences in densities, is called natural convection. Forced convection is achieved by creating fluid motion by an external device, like a pump or fan. The heat transfer rate, \( \dot{Q} \), due to convection can be described by Newton’s law of cooling, according to

\[
\dot{Q} = hA_s (T_s - T_\infty) \quad [W]
\]

where \( A_s \) is the convective area and the temperatures \( T_s \) and \( T_\infty \) are the surface and fluid temperatures, respectively. The parameter \( h \) is the convection heat transfer coefficient and is in the following text referred to as the convection coefficient. Determining the convection coefficient for a certain application is a complex task since it is determined by the combination of several factors, such as the surface geometry, the nature of the fluid, fluid properties and the bulk fluid velocity [24]. Thus, determining the convection coefficient for complex geometries is often done experimentally. Convection coefficients for different cooling methods commonly used for power electronics cooling can be found in the literature and this is further discussed in the section regarding Power Electronic Cooling Methods.

*Radiation* is a mode of heat transfer caused by thermal radiation, which is a type of electromagnetic radiation. Comparing the amount of heat emitted through radiation and the heat removed by forced convection in a power electronic application, shows that the latter is significantly higher. Hence, taking into account that the power devices are limited in temperature and that surfaces with low emissivities are common, the effects of radiation are left out in the thermal analysis carried out in this thesis.

**2.4 Power Electronic Cooling Methods**

Determining the convective heat transfer coefficient for a certain cooling method is often a difficult task, since it involves predicting advanced fluid motions around bodies. Many papers are published, where the convection
coefficients for different cooling methods are determined. The cooling methods can be categorized as follows [25]

- air cooling,
- liquid cooling,
- heat pipes,
- refrigeration cooling,
- thermoelectric cooling,
- phase change material based cooling.

The convection coefficients for the different cooling methods used for power electronic cooling, ranges from approximately 20 W/(m$^2$·°C) up to 300 kW/(m$^2$·°C) [25]-[27]. A range of different convection coefficients is used in the simulations presented in this thesis. In order to cover the wide range of cooling options that could be used in a hybrid vehicle application, convection coefficients ranging from 0,5 kW/(m$^2$·°C) to 50 kW/(m$^2$·°C) are considered. This range covers the cooling methods of forced air cooling via traditional cold plates, to advanced types of liquid cooling, such as spray and jet impingement cooling. The convection coefficients are applied to the bottom of the baseplate or heatsink, depending on the packaging layout or assembly, and are hence assumed to be effective convection coefficients [28].
Chapter 3

Thermal Modelling of Power Modules

3.1 The Heat Conduction Equation

As was pointed out in Chapter 2.3, the main mode of heat transfer occurring in a power module is conduction. The heat generated in the device chips due to the inherent loss generation, is conducted through the layered structure of the device towards the heatsink, see Figure 3.1. Depending on the level of cooling, lateral heat spreading occurs, and this is indicated in Figure 3.1 as a heat spreading angle, \( \nu \).

![Figure 3.1 Heat conduction in power module.](image)

The temperature distribution within the module can be determined by setting up the correct boundary conditions and solving the transient heat conduction equation for each layer within in the module. The transient heat conduction equation for a single layer is given below
where $\lambda$, $\rho$ and $c$ is the thermal conductivity, density and heat capacity, respectively. The generated heat within each layer is denoted $\dot{g}$. Solving the heat conduction equation requires knowledge about the geometry and material composition of the module. In addition to this, initial conditions have to be assumed and boundary conditions have to be set up. For simple geometries and one-dimensional problems, it is quite easy to obtain an analytical solution. However, for complex geometries and if heat spreading and thermal coupling, where chips are heating each other, has to be considered, obtaining a solution is a very complex task. Hence, numerical methods are used to achieve an approximate solution. Studying the literature regarding thermal modelling of power modules, four different methods for obtaining the temperature distribution can be distinguished. These methods are finite difference method (FDM), finite elements method (FEM) and the Fourier series solution method. These numerical methods together with the method of thermal RC-networks are briefly described in the following chapter.

3.2 Solving the Heat Conduction

**Finite Difference Method (FDM)**

The FDM method is based on substituting the derivatives of the governing partial differential equations with differences. To compute these differences, the geometry is divided into a finite number of elements. The accuracy of the final solution is dependent on the number of elements, where more elements will result in a higher accuracy at the expense of more computational effort. The result of the discretization of the geometry and applying that to the heat conduction equation, is a set of approximate algebraic equations. These algebraic equations can then be solved numerically by different methods, such as Jacobi’s method, Gauss-Seidel method or TDMA (Tri-Diagonal-Matrix-Algorithm), to mention a few.

**Finite Element Method (FEM)**

The finite element method is, just as the finite difference method, a numerical method for solving various engineering problems. One of the most important features of the FEM is that it can handle complicated
geometries, whereas the FDM normally is restricted to handle rectangular shapes. As for the FDM, the FEM starts with turning the geometry of interest into small parts or elements. This division of the geometry into sub-parts is commonly referred to as meshing or generating a mesh. The elements are connected at node points. The next step is to approximate the unknown variables with a linear combination of algebraic polynomials and undetermined parameters, where the undetermined parameters are the wanted values of each node [48]. Algebraic relations for the parameters are derived by minimizing a weighted residual for each element. The last step is to put the algebraic relations together based on continuity and “equilibrium”. Solving research and engineering problems with the aid of finite element analysis (FEA) is commonly carried out by using a commercial software package. The FEA tool used in this thesis is called ANSYS Workbench and is developed by ANSYS Inc.

**Fourier Series Solution**

The Fourier series solution method is an analytical solution method which is based on assuming a solution to the heat conduction equation in the form of a Fourier series expansion [29]. This method has in recent years been presented in [30] for a one-dimensional model and later presented in [31] and extended to two dimensional heat conduction problems. The accuracy and the simulation speed of this method depend on the number of harmonics used in the Fourier series expansion [31]. The method is according to [32] faster than the FDM with an acceptable level of accuracy and can easily be programmed in the Matlab/Simulink environment.

**Thermal RC Networks**

If the heat conduction equation is compared to the equation describing the voltage for an electrical transmission line, it can be seen that there are similarities. This equation can, in fact, serve as an electrical analogy of the heat conduction equation after some simplifications. An equivalent circuit diagram for describing a one-dimensional heat flow can be seen in Figure 3.2 Each R/C pair represents a layer in the power module structure. This type of network is often referred to as a Cauer network or model.
In the electrical equivalent, power is equal to current and temperature is equal to voltage. The current source represents the heat dissipated as losses in the device chip. The thermal resistances and capacitances are computed based on material parameters according to

\[
R_{th,i} = \frac{d_i}{\lambda_{th,i} \cdot A_i(v)} \quad [\text{°C/W}]
\]

\[
C_{th,i} = c_i \cdot \rho_i \cdot d_i \cdot A_i(v) \quad [\text{J/°C}]
\]

where \(d_i\) and \(A_i\) are the thickness and area of layer \(i\). The parameters \(\lambda_{th,i}, \rho_i\) and \(c_i\) are the thermal conductivity, density and heat capacity for the material of layer \(i\). Material properties for the materials in the power module are listed in Table 2.1. If a one-dimensional heat flow is assumed, a heat spreading angle, \(v\), can be included to account for heat spreading effects [33]. Convection can also be included in the Cauer model as a thermal resistance. This resistance is obtained by rearranging parts Equation 2.6.

\[
R_h = \frac{1}{h \cdot A_s} \quad [\text{°C/W}]
\]

where \(A_s\) is the convection area. Figure 3.3 shows the thermal resistances calculated based on the dimensions for one of the power modules studied in this project. Different heat spreading angles, \(v\), and convection coefficients, \(h\), ranging from 40° to 70° and 1 to 10 kW/(m²·°C), respectively, are assumed.
3.2. Solving the Heat Conduction

Figure 3.3 Power module thermal resistances (a) for the different layers of an IGBT power module. Thermal resistances due to convection (b).

Figure 3.3(a), shows that the low thermal conductivity of the thermal interface material causes this layer to have the highest thermal resistance. Hence, keeping this layer as thin as possible is important in order to reduce the total thermal resistance. The resistance contribution from the substrate is also quite significant and should thereby be kept as thin as possible, without jeopardizing the electric insulation capacity between the silicon chip and the baseplate. The thermal resistances decreases for a higher heat spreading angle since the assumed active area of each layer increases. The thermal resistance due to convection decreases both with the convection coefficient and the active area, which can clearly be seen in Figure 3.3(b).

Physical based thermal models for power electronics are models that are based on the heat conduction equation. One example is the RC-network of the Cauer type, where the parameters are computed based on the geometrical and material properties of the power module structure. The Foster network is, in contrast to the Cauer network, not a physical based model. In fact it is often referred to as being a behavioural model. The reason for this originates from the fact that the thermal model is derived by fitting functions to measurements or simulation results obtained from FEA.
calculations. An equivalent RC-network of the Foster type can be seen in
Figure 3.4.

\[ Z_{th} = \frac{T_j - T_a}{P} \quad [{\degree C/W}] \quad (3.5) \]
3.3. Transient Thermal Impedances

The thermal impedance is actually the unit step response. Three thermal impedances for an IGBT power module with three different convection coefficients are shown in Figure 3.5.

![Figure 3.5 Thermal impedance for an IGBT with three different convection coefficients.](image)

The figure shows that for a pulse length of less than 1 s, the thermal impedance of the device is independent on the level of convection. The reason for this is the lower thermal capacitances of the chip and the substrate compared to the baseplate. Hence, all of the dissipated heat is, during the duration of the pulse, stored in the chip and substrate. For pulse length above 1s, the heat has transferred down into the baseplate or heatsink. Hence, having a lower convection coefficient results in higher thermal impedance. The thermal impedance value for infinitely long pulses is equal to the steady state thermal resistance.

The transient thermal impedance for a single-pulse, presented in Figure 3.5, is more relevant for application as welding or induction heating than electrical drive applications. Hence, transient thermal impedance curves are often given for repetitive pulses of different duty cycles. Figure 3.6
shows the transient thermal impedance for repetitive pulses of different pulse widths and duty cycles ranging from 0 to 0.5.

Figure 3.6 Thermal impedance for five different duty cycles with a convection coefficient of 5 kW/(m²·°C).

The heat losses generated in power semiconductor used in drive applications are normally of a more repetitive nature. In a three-phase inverter used for controlling an electrical machine, the diodes and switches of a phase leg are operated in pairs. During a switching period, depending on the current direction, it is either the upper transistor and the lower diode or the lower switch and upper diode, that are operated together to conduct the phase current. Hence, during one electrical period the switch/diode pairs will be active 50% of the period. Assuming electrical frequencies ranging from 0-500 Hz, the thermal impedance curves for pulse widths ranging from 5 ms to infinity with a duty cycle of 50% are the most relevant for power devices used in drive applications. Figure 3.6 shows that the maximum junction temperature increases as the duty cycle is increased, since an increased duty cycle implies less time for the chip to cool off.
3.4 Mathematical Derivation of Foster Based Thermal Models

The development of the Foster network has its origin in Duhamel’s theorem. This theorem is an analytical method that can be used to solve transient heat conduction problems when time-dependent functions are present. These functions can either appear in the boundary conditions and/or as source terms [37]. Duhamel’s theorem can be written as

\[
T(r, t) = T_0 + \int_0^t f(\tau) \left[ \frac{\partial u(r, t-\tau)}{\partial t} \right] d\tau
\]  

for the temperature evolution at point \( r \). The time dependent boundary condition is denoted \( f(t) \) and the function \( u(t) \) is the temperature response to a unit step. The theorem can be used to compute the junction temperature response to its internal power loss. It can also be extended to include the internal heating or thermal coupling between devices taking place inside multi-chip power modules. If only a single heat source is considered, the temperature, \( T \), at time \( t \) at point \( r \) can be computed by solving the Duhamel integral

\[
T(r, t) = T_0 + \int_0^t P_i(\tau) \frac{\partial Z_{th}(r, t-\tau)}{\partial t} d\tau
\]  

where \( T_0 \) is the initial temperature. The time dependent function \( f(t) \) and the unit step response \( u(t) \) in Equation 3.6 is, when applying the theorem to the calculation of the power module junction temperature, equal to the heat generated in the chip, \( P(t) \), and the thermal impedance, \( Z_{th}(t) \). As will be studied later, the thermal impedance can be determined by fitting a suitable function to measurements or the results from FEA calculations. For a multichip module, more than one heat source needs to be taken into account. This can be done by applying the superposition principle [38]. The temperature evolution of device \( j \), \( T_j(t) \), due to several heat sources \( P_i \) can be written as

\[
T_j(t) = T_0 + \sum_{i=1}^{N} \int_0^t P_i(\tau) \frac{\partial Z_{th,i,j}(t-\tau)}{\partial t} d\tau
\]
where $N_i$ is the total number of heat sources. The thermal impedance $Z_{th,ij}$ is the thermal response of chip $j$ to a unit power step applied to chip $i$. The space dependence on $r$ has been dropped in the above equation, since it is assumed that all temperatures are calculated at the centre of every chip. The thermal impedance matrix, including thermal coupling between the different devices within the power module, can be written as

$$Z_{th} = \begin{bmatrix}
Z_{th,11} & \cdots & Z_{th,1j} & \cdots & Z_{th,1n} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
Z_{th,n1} & \cdots & Z_{th,nj} & \cdots & Z_{th,nn}
\end{bmatrix} \quad (3.9)$$

Each element of the transient thermal impedances matrix given in Equation 3.9 is determined from FEA simulations or measurements and can be computed according to

$$Z_{th,ij}(t) = \frac{T_j(t) - T_a}{P_i(t)} \quad (3.10)$$

where $T_a$ is the ambient temperature, here represented by the temperature of the cooling medium.

An analytical expression of the thermal impedances can be derived by fitting a Foster network to the impedance curves obtained from FEA simulations or measurements. The fitted function used here is a weighted sum of $N_k$ exponentials [36].

$$Z_{th,ij}(t) = \sum_{k=1}^{N_k} R_{th,ij}^k \cdot \frac{1 - e^{-\frac{t}{\tau^k_{th,ij}}}}{\tau^k_{th,ij}}, \quad \tau^k_{th,ij} = R_{th,ij}^k \cdot C_{th,ij}^k \quad (3.11)$$

The $R_{th,ij}^k$ and $C_{th,ij}^k$ elements are the resistive and capacitive electrical equivalents of the Foster network. The number of exponentials, $N_k$, that is needed to accurately represent the FEA calculations depends on whether the impedance is a mutual ($i \neq j$) or self ($i = j$) impedance. It is found that no
more than 6 exponentials are needed for the self-impedances, whereas only 2 are needed for the mutual. Hence, if the temperatures of the different layers within the devices are of no interest, there is no need for using more than 6 exponentials. However, if the different layer temperatures are of interest, two exponentials per device layer should be used, resulting in 14 exponentials for an ordinary device [40]. Figure 3.7 shows a schematic overview of a typical RC-network of Foster type. The figure shows how the junction temperatures of the two diodes of a phase leg (refer to Figure 4.4 in Chapter 4.5) are modelled.

Figure 3.7  Schematic overview of a RC-network of Foster type.

Included in the figure are the mutual impedances (cross-coupling), representing the thermal interdependencies between the devices, and the self-impedances, representing the heating of the device itself. The current sources represent the power losses generated in each device. A voltage source is added to represent the temperature of the cooling medium.
The convolution integral, given in Equation 3.7, can be solved using several different methods. One way is to solve the integral in the frequency domain by using the fast Fourier transform of the derivative of $Z_{th,ij}$ and $P(t)$ [36]. The temperature evolution of device $j$ can then be computed as

$$T_j(t) - T_0 = \sum_{i=1}^{N} FFT^{-1}\{Z_{th,ij}(\omega) \cdot P_j(\omega)\}$$

(3.12)

The drawback of this method is the calculation speed, since the Fourier transform needs to be calculated at each time step. Another method to compute the convolution integral is to solve it numerically by a recurrent algorithm as presented in [39]. This requires analytical expressions of the thermal impedances, preferably in the form of exponential functions. This method has been proven to be very accurate and have high performance in terms of computation time. The integral can also be solved by using the Laplace transform to turn the model into a state space equation. The temperature increase of chip $j$ due to a power loss in chip $i$ can, in state space form, be written as

$$\Delta T_{ij}(s) = A_{ij} \cdot \Delta T_{ij}(s) + B_{ij} \Delta P(s)$$

$$\Delta T_j(s) = C_{ij} \Delta T_j(s)$$

(3.13)

where the state vector and its time derivative are defined as

$$\Delta \bar{T}_j(s) = \begin{bmatrix} \Delta T_{i1}(s) \\ \vdots \\ \Delta T_{in}(s) \end{bmatrix}, \Delta \tilde{T}_j(s) = \begin{bmatrix} \Delta T_{ij}(s) \\ \vdots \\ \Delta T_{in}(s) \end{bmatrix}$$

(3.14)

The definition of the matrices $A_{ij}$, $B_{ij}$ and $C_{ij}$ can be found in Appendix A. The total temperature evolution of chip $j$ can then be written as

$$\Delta \bar{T}_j(s) = A_j \cdot \Delta \bar{T}_j(s) + B_j \Delta \tilde{T}(s)$$

$$T_j(s) = C_j \Delta \bar{T}_j(s) + T_a$$

(3.15)

where the state vector and its time derivative are defined as
3.4. Mathematical Derivation of Foster Based Thermal Models

\[ \Delta \dot{T}_j(s) = \begin{bmatrix} \Delta \dot{T}_{i,j}(s) \\ \vdots \\ \Delta \dot{T}_{n,j}(s) \end{bmatrix}, \Delta \bar{T}_j(s) = \begin{bmatrix} \Delta T_{i,j}(s) \\ \vdots \\ \Delta T_{n,j}(s) \end{bmatrix} \]  

(3.16)

The input vector is defined as

\[ \bar{P}(s) = \begin{bmatrix} P_i(s) \\ \vdots \\ P_n(s) \end{bmatrix} \]  

(3.17)

The definition of the matrices \( A_j, B_j \) and \( C_j \) can be found in Appendix A. The solution method used in this thesis is the Laplace transform method, mainly due to its ease of derivation and suitability to be implemented and solved in Matlab/Simulink.
Chapter 4

System Modelling

4.1 System Model

The simulation model used in this thesis describes the EDS of a hybrid car. The model is divided into three main parts, the vehicle model, the electrical model and the thermal model, see Figure 4.1.

![Figure 4.1 Full system simulation model overview.](image)

The output from the vehicle model is the requested electrical machine torque and speed. The torque and speed combinations are computed based on a number of well-known driving cycles and a simple longitudinal model of the vehicle. The electrical model consists of two sub models, one for the electrical machine and one for the power electronic losses, and a transformation block. The output of the electrical model is the inverter losses, which are used as inputs to the thermal model. This model contains...
Chapter 4. System Modelling

the thermal models for a number of different power modules, and computes the junction temperatures of the power devices. Each of these parts is described more in detail in the subsequent sections.

4.2 Power Modules

Five different power modules are studied in the simulations carried out in this thesis. Module data in terms of included materials, the geometrical layout and electrical parameters are presented in Table 4.2. Schematic overviews of the module layouts, together with device naming, are presented in Figure D.1 to Figure D.4 in Appendix A. Each of the power modules is assigned a number and a code for the type of cooling. The codes for the different cooling options are given in Table 4.1.

Table 4.1 Abbreviations for the different power module cooling options.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>Baseplate cooling</td>
</tr>
<tr>
<td>h</td>
<td>Module baseplate is mounted on a heatsink</td>
</tr>
<tr>
<td>hnb</td>
<td>Module does not have a baseplate. Module substrate is mounted directly on a heatsink.</td>
</tr>
<tr>
<td>dsc</td>
<td>Module does not have any heatspreader (baseplate or heatsink). Cooling is applied both to the top and bottom of the module.</td>
</tr>
</tbody>
</table>

The power modules are in the following text referred to with this specific number in order to simplify for a direct comparison of the different modules. As an example, power module number 1 with baseplate cooling is referred to as PM1b.

The power electronics used within the eTVD project uses a 400A three-phase power module. The reason for studying a single-phase module, in addition to the three-phase module, is that it is assumed that some thermal and packaging benefits are gained. The reason for studying modules with two different current levels, i.e. 300A and 400A, is that the SiC BJT module studied here is under development and is currently, only manufactured in a 300A-version. Thus, to make a fair comparison, a 300A single-phase module is added to the study.
### 4.2. Power Modules

Table 4.2  Power module data.

<table>
<thead>
<tr>
<th>Assembly</th>
<th>PM1</th>
<th>PM3</th>
<th>PM4</th>
<th>PM5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Module layout</strong></td>
<td>Figure D.1</td>
<td>Figure D.2</td>
<td>Figure D.3</td>
<td>Figure D.4</td>
</tr>
<tr>
<td><strong>Module type</strong></td>
<td>3-phase</td>
<td>1-phase</td>
<td>1-phase</td>
<td>1-phase</td>
</tr>
<tr>
<td><strong>Switch type</strong></td>
<td>IGBT</td>
<td>BJT</td>
<td>IGBT</td>
<td>IGBT</td>
</tr>
<tr>
<td><strong>Nr of dev. per switch</strong></td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Nr of dev. per diode</strong></td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Electrical parameters

<table>
<thead>
<tr>
<th></th>
<th>Vce [V]</th>
<th>Ic,nom [A]</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM1</td>
<td>650</td>
<td>400</td>
</tr>
<tr>
<td>PM3</td>
<td>1200</td>
<td>300</td>
</tr>
<tr>
<td>PM4</td>
<td>600</td>
<td>300</td>
</tr>
<tr>
<td>PM5</td>
<td>600</td>
<td>400</td>
</tr>
</tbody>
</table>

#### Geometrical parameters

<table>
<thead>
<tr>
<th>Chip switch</th>
<th>Si</th>
<th>SiC</th>
<th>Si</th>
<th>Si</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,2x9,7x0,07</td>
<td>5,4x3,5x0,40</td>
<td>10,5x10,5x0,10</td>
<td>11,6x11,6x0,10</td>
<td></td>
</tr>
<tr>
<td>Chip diode</td>
<td>Si</td>
<td>SiC</td>
<td>Si</td>
<td>Si</td>
</tr>
<tr>
<td>5,4x9,2x0,07</td>
<td>8,3x2x0,40</td>
<td>7,8x7,0x0,35</td>
<td>9,0x8,6x0,35</td>
<td></td>
</tr>
<tr>
<td>Solder switch</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>10,2x9,7x0,1</td>
<td>5,4x3,5x0,05</td>
<td>10,5x10,5x0,10</td>
<td>11,6x11,6x0,10</td>
<td></td>
</tr>
<tr>
<td>Solder diode</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5,4x9,2x0,1</td>
<td>8,3x2x0,05</td>
<td>7,8x7,0x0,10</td>
<td>9,0x8,6x0,10</td>
<td></td>
</tr>
<tr>
<td><strong>DBC</strong></td>
<td>Cu/Al2SO3/Cu</td>
<td>Cu/Al2SO3/Cu</td>
<td>Cu/Al2SO3/Cu</td>
<td>Cu/Al2SO3/Cu</td>
</tr>
<tr>
<td>37,2x57,0x0,92</td>
<td>57,8x40,8x0,98</td>
<td>63,0x37,0x0,75</td>
<td>63,0x37,0x0,75</td>
<td></td>
</tr>
<tr>
<td>Solder</td>
<td>-</td>
<td>x</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>37,2x57,0x0,20</td>
<td>x</td>
<td>63,0x37,0x0,3</td>
<td>63,0x37,0x0,3</td>
<td></td>
</tr>
<tr>
<td>Baseplate</td>
<td>Cu</td>
<td>x</td>
<td>Cu</td>
<td>Cu</td>
</tr>
<tr>
<td>140x72x3</td>
<td>x</td>
<td>90x43x3</td>
<td>90x43x3</td>
<td></td>
</tr>
<tr>
<td><strong>Thermal grease</strong></td>
<td>Silicone paste</td>
<td>Silicone paste</td>
<td>Silicone paste</td>
<td>Silicone paste</td>
</tr>
<tr>
<td>140x72x0,04</td>
<td>57,8x40,8x0,04</td>
<td>90x43x0,04</td>
<td>90x43x0,04</td>
<td></td>
</tr>
<tr>
<td><strong>Heatsink</strong></td>
<td>Aluminium</td>
<td>Aluminium</td>
<td>Aluminium</td>
<td>Aluminium</td>
</tr>
<tr>
<td>210x142x5</td>
<td>128x110x5</td>
<td>160x130x5</td>
<td>160x130x5</td>
<td></td>
</tr>
</tbody>
</table>
4.3 Vehicle Model

The vehicle studied in this thesis is, as has been mentioned before, a PHEV. A PHEV has, apart from the ICE, at least one additional source of power available for vehicle propulsion. The control of these power sources is managed by an energy management controller and its associated energy control strategy. The control strategy has several different objectives, where the primary ones usually aim to minimize the fuel consumption and emissions. The control objectives are achieved by instantaneous management of the power split between the power sources. This power split is crucial for the overall performance of the HEV system. Apart from achieving a high fuel economy and minimizing the emissions, the strategy must maintain the drivability or preferably enhance it as is the case with the eTVD-system. The basic energy management strategy for a PHEV is to run the vehicle in pure electric mode as much as possible and only turn on the ICE if the power requirement for the EDS is too high or the state of charge of the battery is too low. If the ICE is turned on, the energy management strategy is activated and controls the power flow to the ICE and the EM in order to minimize the fuel consumption. However, the most demanding type of strategy from an EDS perspective would be to supply as much power as possible with the EDS and only use the ICE to supply the shortage. The reason for assuming an energy management strategy implemented according to this, is to simplify the simulation model and to create a worst case scenario for the EDS.

The model used for the vehicle is a rather simple longitudinal model. Hence, vehicle stability and forces acting in other directions are not considered. The force needed to achieve an acceleration \( a \) of a vehicle with a mass \( M_v \) at a certain velocity \( v \), taking into account rolling friction, air drag and inertial forces, is given by Equation 4.1

\[
F_{\text{veh}} = M_v a + C_r(v) M_v g + \frac{1}{2} \rho_a C_d A_v v^2 + F_I
\]  

(4.1)

where \( C_r \) is the rolling resistance coefficient, \( \rho_a \) is the air density, \( C_d \) is the aerodynamic drag coefficient, \( A_v \) is the vehicle front area, \( g \) is the acceleration of gravity and \( F_I \) the inertial forces.

The inertial forces are a function of the vehicle acceleration. The forces
associated with the acceleration of the different inertias can be computed according to Equation 4.2

$$F_j = \left(C_{\text{axle},f} + C_{\text{axle},r} + \left(C_{\text{EM}} + C_{g}\right) \cdot r_{\text{EM}}^2\right) \cdot \frac{dv}{dt}$$  (4.2)

where the different inertias and gear ratios are defined according to Table 4.3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$J_{\text{axle},f}$</td>
<td>Front axle inertia</td>
</tr>
<tr>
<td>$J_{\text{axle},r}$</td>
<td>Rear axle inertia</td>
</tr>
<tr>
<td>$J_{\text{EM}}$</td>
<td>EM inertia</td>
</tr>
<tr>
<td>$J_{g}$</td>
<td>Total inertia of gear box</td>
</tr>
<tr>
<td>$r_{\text{EM}}$</td>
<td>EM gear ratio</td>
</tr>
</tbody>
</table>

Data for all the vehicle parameters are given in Appendix C.

### 4.4 Electrical Machine Model

The electrical machine used in the hybrid vehicle is of the type interior permanent magnet synchronous machines (IPMSM). The reason for using this type of machine is that it has high efficiency and high power density, requires low maintenance and is relatively easy to control [45]. All this together makes the PMSM a competitive choice when it comes to tractive applications. However, it has its drawbacks, such as relatively low efficiency for high speeds, due to the demagnetizing current needed for field weakening, and having expensive permanent magnets as the source of magnetic flux. Another drawback is that it has a low speed ratio, i.e. low ratio between the maximum speed and the base speed, which is disadvantageous for vehicle acceleration compared to electrical machine power. The EM torque and power as a function of speed can be seen in Figure 4.2(b) and all basic data for the EM is given in Table 4.4.
Figure 4.2  EM motoring torque as a function of d- and q- currents, and the MTPC current combinations (a). EM peak power and peak torque (b).

Table 4.4    Electrical machine data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cont./Peak power, $P_{EM,cont}/P_{EM,peak}$</td>
<td>67/93</td>
<td>kW</td>
</tr>
<tr>
<td>Cont./Peak torque, $T_{EM,cont}/T_{EM,peak}$</td>
<td>225/330</td>
<td>Nm</td>
</tr>
<tr>
<td>RMS-current cont/peak, $I_{cont}/I_{peak}$</td>
<td>200/300</td>
<td>A</td>
</tr>
<tr>
<td>Max. operating speed, $n_{EM,max}$</td>
<td>10500</td>
<td>rpm</td>
</tr>
<tr>
<td>Base speed, $n_{base} (U_{dc}=320V)$</td>
<td>2500</td>
<td>rpm</td>
</tr>
</tbody>
</table>

The system equations for an IPMSM are given in Equation 4.3 and 4.4 below

$$\frac{d\psi_d(i_d,i_q)}{dt} = u_d - R_s i_d + \omega_e \psi_q(i_d,i_q)$$  \hspace{1cm} (4.3)
where $\psi_d$, $\psi_q$, $i_d$, $i_q$, $u_d$, and $u_q$ are the flux, current, and voltage in the d- and q-direction, respectively. The stator winding resistance is denoted $R_s$ and $\omega_e$ is the electrical machine frequency. The mechanical torque, $T_m$, of the EM can be computed according to

$$ T_m = \frac{3n_p}{2K^2} \left( \psi_d \left( i_d, i_q \right) \cdot i_q - \psi_q \left( i_d, i_q \right) \cdot i_d \right) $$

where $n_p$ are the number of pole pairs and $K$ is a parameter related to the type of transformation used. The amplitude invariant transformation is used throughout the thesis, and $K=1$ for this transformation. The flux look-up tables, $\psi_d(i_d,i_q)$ and $\psi_q(i_d,i_q)$, are derived from FEM calculations and by using these as a base for the EM model, saturation and saliency effects, such as inductance variations and reluctance torque are included in the model. The input to the EM-model is an operating point defined by the EM rotational speed and the requested torque, both determined by the driving cycle. Since the three phase voltages and currents of the machine are needed in order to compute the power electronic losses, each operational point has to be converted to a voltage and current combination. An algorithm that uses the FEM results, computes the current combination that produces the requested torque \cite{41}. Several current combinations can produce the same torque and the combination that maximizes the torque per current ratio (MTPC) is chosen, see Figure 4.2(a). The algorithm also takes current and voltage limits into account, so that the torque is maximized as the speed of the machine is increased. The output of the algorithm is look-up tables for the $i_d$ and $i_q$ current as a function of torque and speed, see Figure 4.3 below.
Figure 4.3 Currents in the rotor reference frame, $i_d$ (a) and $i_q$ (b), as a function of the torque request and rotational speed.

The current look-up tables are in combination with Equation 4.3 and 4.4 used to compute the voltages. The derivatives in these equations are set to zero, resulting in a steady state model. The motivation for using a steady-state model is that the simulation model is greatly simplified, since no current controller has to be implemented. Using a steady state model does not introduce any significant error since the requested torque is rate limited in order to protect the mechanical construction of the rear-wheel drive. A rate limit on the order of 5000 Nm/s results in a rise time of approximately 60ms for the full torque range. Assuming that the current rise time is the same and a winding inductance in the range of 0.2–1 mH, results in an inductive voltage drop of 1–5V, which is considered to be negligible. These assumptions have been verified by comparing a transient simulation of the EDS with the steady state model.

4.5 Power Electronic Loss Modelling

Real-time simulation of switching waveforms for IGBT:s and diodes requires very short time-steps, due to the normally short switching times
4.5. Power Electronic Loss Modelling

(in the range of nano-seconds [42]) and high dynamics. Studying the thermal response for mission profiles lasting several hundreds of seconds, would result in unrealistically long simulation times. One solution, which also is used in [42]-[44], is to compute the average loss for the semiconductors, i.e. IGBT and diode, over one switching period. This, together with the fact that the smallest thermal time constant for the power module is in the range of 1 ms, which will be shown in a later section, implies that a simulation step in the range of the switching period should be used. Averaging of the inverter losses makes it possible to combine real-time simulation of device losses with thermal simulations. The device losses are computed for a range of device temperatures, load currents and duty ratios, and are stored in look-up tables. The average losses for the IGBT:s and the diodes over one switching period are computed from data taken from the datasheet of the device. The key parameters for computing the IGBT and the diode losses are the turn-on, $E_{T,\text{on}}$ and $E_{D,\text{on}}$, and turn-off, $E_{T,\text{off}}$ and $E_{D,\text{off}}$, energies, together with the on-state voltage drops, $V_{ce}$ and $V_{f}$. These parameters are given as a function of the device current, $I_L$, and device temperatures, $T_{Tj}$ and $T_{Dj}$, respectively. The energy dissipation originating from the conduction losses, $E_{T,c}$ and $E_{D,c}$, over one switching period, $T_{sw}$, for the IGBT:s and diodes, can be calculated as

$$E_{T,c} = \int_{t_{\text{T,on}}}^{t_{\text{T,off}}} P_{T,c}(t) \, dt = \frac{d_f V_{ce} |I_L|}{f_{sw}}$$  \hspace{1cm} (4.6)

$$E_{D,c} = \int_{t_{\text{D,on}}}^{t_{\text{D,off}}} P_{D,c}(t) \, dt = \frac{d_B V_{f} |I_L|}{f_{sw}}$$  \hspace{1cm} (4.7)

where $t_{\text{T,on}}$ and $t_{\text{D,on}}$ are the on-times of the IGBT and diode in an IGBT-diode pair, respectively. The duty cycles for the IGBT:s and the diodes are denoted, $d_f$ and $d_B$, respectively. It is assumed that the current and consequently the voltage, are constant during the on time of the device. Hence, the current ripple, originating from the switching, is not assumed to have any significant impact on the module from a thermal perspective. The EM voltages and currents are transformed from synchronous coordinates to three phase quantities in the transformation block, see Figure 4.1. The transformation block also computes a voltage ratio, $d$, which, for all three phases of the inverter, is defined as
\[ d_x = \frac{v_x^*}{U_{dc}/2} \]  

(4.8)

where \( x \) denotes phase \( a, b \) or \( c \) and \( v_x^* \) is the instantaneous phase voltage reference. Symmetrized voltage references are used in order to make full use of the DC-voltage. The voltage ratio is used in the inverter model to compute the duty cycles for the semiconductor devices in a phase leg. Figure 4.4 shows an inverter phase leg.

![Inverter phase leg](image)

Figure 4.4 Inverter phase leg.

The duty cycles, \( d_I \) and \( d_D \), for each of the devices shown in Figure 4.4 are defined according to Table 4.5.

<table>
<thead>
<tr>
<th>Current direction</th>
<th>IGBT 1 (upper)</th>
<th>IGBT 2 (lower)</th>
<th>Diode 1 (upper)</th>
<th>Diode 2 (lower)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( I_L &lt; 0 )</td>
<td>-</td>
<td>( d_I = (1-d_D)/2 )</td>
<td>( d_D = (1+d_I)/2 )</td>
<td>-</td>
</tr>
<tr>
<td>( I_L &gt; 0 )</td>
<td>( d_I = (1+d_D)/2 )</td>
<td>-</td>
<td>-</td>
<td>( d_D = (1-d_I)/2 )</td>
</tr>
</tbody>
</table>

Table 4.5 Device duty cycle definition.

The total average power losses over one switching period can finally be computed according to Equation 4.9 and 4.10 and stored as look-up tables.

\[
P_{T,\text{loss}}(d_T, I_L, T_j) = f_{sw} \cdot (E_{T,c} + E_{T,on} + E_{T,off})
\]  

(4.9)

\[
P_{D,\text{loss}}(d_D, I_L, T_D) = f_{sw} \cdot (E_{D,c} + E_{D,on} + E_{D,off})
\]  

(4.10)
The total power losses as a function of load current and duty cycle, for a junction temperature of 125 °C and switching frequency of 10 kHz, for the IGBT and diode of PM5 are shown in Figure 4.5(a-b).

Figure 4.5 IGBT (a) and diode (b) losses for PM5 as a function of load current and duty cycle at a junction temperature of 125 °C and a switching frequency of 10 kHz. IGBT (c) at a junction temperature of 125 °C and a switching frequency of 3 kHz. Decrease in total IGBT losses (d) when the switching frequency is reduced from 10 to 3 kHz.

Figure 4.5(c) also shows the total power losses for the IGBT when a switching frequency of 3 kHz is used. It can be seen that the losses are reduced significantly, and Figure 4.5(d) shows the total reduction ranges between 20 to 70% depending on the operating point. Using a variable switching frequency to reduce the total losses is discussed in Chapter 5.4.

The losses for both of the devices are increasing linearly with the duty cycle, whereas the losses are increasing exponentially with the load current. The loss maps for the other power modules have similar shapes, and it is mainly the absolute values of the losses that are different.
Chapter 4. System Modelling

Computing the losses based on datasheet values and using the equations presented above, gives the total amount of losses for each device, i.e., for each IGBT or diode. However, the modules studied in this thesis contain several parallel connected chips for each device. The total losses are simply split equally for each chip that is connected in parallel. This is motivated by the fact that the thermal models, presented in Chapter 4.6, indicates that the difference in junction temperatures of the parallel connected chips are low.

The battery voltage is assumed to be constant and a fixed DC-voltage of 320 V is used in all simulations presented in the thesis.

4.6 Thermal Models of Power Modules

Thermal Model Development Procedure

The thermal models for the power modules studied in this thesis are all derived according to the same procedure. An overview of the procedure is presented in Figure 4.6.

The first step is to gather geometrical and material data for each of the layers within the module. These data are collected and transformed into a file that can be read by the FEA-software and used to set-up a finite element model of the module.
In the second step, the model is run as a batch where a fixed amount of heat is injected into each device separately. The junction temperature evolutions at the centre of all chips are recorded and saved to a file. This is repeated for several convection coefficients, ranging from 0.5 to 50 kW/(m²·°C). Constant convection coefficients are applied to the lower surface of the bottom layer, which can be either the baseplate or heatsink depending on how the module is integrated. As has been addressed before, computing the convection coefficient for a specific cooling method is a difficult and a computational burdensome task and a detailed analysis regarding the convective mechanism is out of scope for this thesis. Instead, a simplified approach is adapted where a constant effective convection coefficient and a constant cooling liquid temperature is assumed. The results from the FEM calculations are thermal impedances describing the thermal interaction between the chips within the power module. These thermal impedances are computed according to Equation 3.5, repeated here for clarity,

\[ Z_{th,ij}(t) = \frac{T_j(t) - T_a}{P_i(t)} \]  

where \( Z_{th,ij} \) is the thermal impedance describing the temperature rise of chip nr \( j \) due to heat injected in chip nr \( i \). The results from the FEA, in the form of thermal impedance data, are sent to a curve fitting algorithm.

In the third step, an algorithm, based on a nonlinear least square procedure, fits Foster networks, see Equation 3.11, to the thermal impedance data obtained from the FEA. The number of exponentials used is 4-5 for the self-impedances and 1-2 for the mutual. The foster network parameters, in the form of thermal resistances and capacitances, are used in the next step to derive a state space model of the thermal system. The state space model is derived according to the procedure described in Chapter 3.4. Finally, the state space models are incorporated in the full system simulation model.

**Thermal Model Analysis**

The impedances obtained from the FEA and the errors compared to fitted foster networks for PM4h, when applying heat to chip nr 5 (T11), is shown in Figure 4.7.
Figure 4.7 Self and mutual impedances for a convection coefficient of 0.5 (a) and 5 kW/(m$^2$·°C) (c), when heat is dissipated in chip T11 of PM4h. Difference in thermal impedance between FEM results and fitted foster networks (b and d).
Figure 4.7 shows that there is a thermal coupling between the chips within the module. This coupling originates from the heat spreading within the power module and is important to incorporate into the thermal model. Changing the convection coefficient changes the strength of the coupling, which is exemplified by the results presented in Figure 4.7 where the steady state values of the mutual thermal impedances are significantly reduced as the convection coefficient is increased. Increasing the convection coefficient lowers the thermal impedance of the cooler itself, resulting in lower lateral temperature gradients and thereby less lateral heat spreading. Figure 4.7 shows that the temperature starts to rise immediately in the chip where the heat is generated. The temperature of the other chips rises after a delay which is related to the distance from the heat source, since a finite amount of time is required for the heat to propagate within the module. Omitting the thermal coupling within the module leads to large model errors. Figure 4.8(a) and (b) show the temperature rise for one of the IGBT chips (T11) of PM1 and PM5, when a loss of 100W is assumed for each of the devices of the upper IGBT and the lower diode for one phase leg. The temperature rise is computed for two different models, where one consists of a single thermal impedance for each chip and the other one, which in addition to this, takes the thermal coupling into account.
Figure 4.8 Temperature rise of chip T11 for PM1b (a) and PM5b (b), when 100 W is injected in each of the devices of T1 and D2, for convection coefficients ranging from 1.0 to 30 kW/(m²·°C).
The figures show that the error in the predicted temperature rise, when comparing the two models, is very high, around 70 °C, when the convection coefficient is low, but almost negligible, around 2 °C, for high convection coefficients. Hence, for convection coefficients lower than 30 kW/(m²°C), the thermal coupling within the module needs to be taken into account to avoid underestimating the chip temperatures.

The differences in the thermal impedances between the results obtained from the FEA and the fitted Foster network, when thermal coupling is taken into account, are presented in Figure 4.7(b and c). The differences are small and the maximum deviation is around 0.005 °C/W, which would result in a steady state temperature error of 1 °C for a loss power of 200 W. This accuracy is considered to be sufficient for the use in the full system simulation model.

Three different module assemblies with single-sided cooling are studied in this thesis. In the first assembly, the power module is mounted on a 5mm thick aluminium heatsink and a thermal compound is used between the power module baseplate and the heatsink to assure good thermal contact. The second assembly is simply the power module itself, where the module baseplate serves as the thermal heatsink. In the third assembly, the baseplate is omitted and the module substrate is mounted directly on the heatsink via a thermal compound. Apart from the three different single-sided cooled module assemblies described above, one assembly with double-sided cooling is studied. The type of layout and assembly greatly affects the thermal impedance of the power module, which can be seen in the figures presented in the subsequent sections. What impact the different layouts and assemblies have on the thermal impedance is further discussed in Chapter 5.5.

**Thermal Models for the Different Power Modules**

The self and mutual thermal impedances for all the modules studied in this thesis are presented in Figure A.1 to Figure A.8, in Appendix A. Thermal impedances for different module assemblies and convection coefficients are presented.
Chapter 5

System Simulations

5.1 System Simulations and Evaluation Criteria

Several system simulations are carried out in order to determine the cooling requirements and its effect on module reliability. To determine the cooling requirements, some kind of criteria has to be defined. The criterion used in this thesis is that the junction temperature of all devices within a module should be kept below 125°C. This is a rather simple criterion and does not take into account the actual temperature swings and the mean temperatures of the devices, which are important to consider from a reliability perspective. Hence, in addition to the maximum temperature criteria, a study of the reliability using the methods presented in Chapter 2.2 is carried out in order to give a more complete input to the process of determining the cooling requirements.

In addition to different driving cycles, the system is evaluated for different convection coefficients, module assemblies and cooling medium temperatures, together with a comparison of using a fixed and variable switching frequency. All of these different variations to the system are described in the subsequent chapters.

5.2 Driving Cycles

The main part of the different driving cycles are well known cycles, frequently used for hybrid simulations. The cycles have different properties in terms of cycle length, maximum and mean speed and deceleration/acceleration. All of the data is given in Table 5.1.
Table 5.1 Driving cycle properties.

<table>
<thead>
<tr>
<th>Driving cycle</th>
<th>Cycle length [s]</th>
<th>Max. speed [km/h]</th>
<th>Mean speed [km/h]</th>
<th>Max. dec./acc. [m/s²]</th>
<th>Mean. dec./acc. [m/s²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEDC</td>
<td>1190</td>
<td>120</td>
<td>46</td>
<td>-1.6/1.1</td>
<td>-0.8/0.6</td>
</tr>
<tr>
<td>US06</td>
<td>596</td>
<td>129</td>
<td>83</td>
<td>-3.1/3.8</td>
<td>-0.7/0.7</td>
</tr>
<tr>
<td>HYZEM (urban)</td>
<td>559</td>
<td>57</td>
<td>29</td>
<td>-2.0/2.0</td>
<td>-0.6/0.6</td>
</tr>
<tr>
<td>HYZEM (rural)</td>
<td>843</td>
<td>103</td>
<td>53</td>
<td>-3.8/2.2</td>
<td>-0.6/0.5</td>
</tr>
<tr>
<td>HYZEM (highway)</td>
<td>1804</td>
<td>138</td>
<td>95</td>
<td>-3.9/2.0</td>
<td>-0.4/0.3</td>
</tr>
<tr>
<td>FTP75</td>
<td>1375</td>
<td>204</td>
<td>85</td>
<td>-3.6/3.4</td>
<td>-1.3/1.1</td>
</tr>
</tbody>
</table>

5.3 Convection Coefficients

The range of different convection coefficients used in the simulations is intended to reflect the wide range of power electronic cooling methods. As has been addressed before, computing the convection coefficient for a specific cooling method is a difficult and a computational burdensome task and a detailed analysis regarding the convective mechanism is out of scope for this thesis. Instead, a simplified approach is adapted were a constant effective convection coefficient and a constant cooling liquid temperature is assumed.

5.4 Fixed and Variable Switching Frequency

The switching power losses in the power devices are, according to Equation 4.9 and 4.10, proportional to the switching frequency. The ratio between the switching losses and the conduction losses is dependent on the operating point of the converter. If there is no load current, the total power losses will only consist of switching losses, whereas if the load current and duty cycle is high, the conduction losses will dominate. From a thermal and reliability point of view, it is obvious that the switching frequency should be kept as low as possible to minimize the losses and thereby limit the maximum junction temperature. However, the switching frequency cannot be chosen arbitrarily low, since this will degrade the performance of the drive system. A low switching frequency will increase the phase current ripple, which in turn will have a negative impact on the EM torque
5.5. Power Module Integration

quality. An EM with high inductance has lower current ripple and is less sensitive to switching frequency reductions. Reducing the switching frequency from 10 kHz to 3 kHz for the machine used in this thesis, increases the torque ripple from 10 Nm to 20 Nm for a current of 70% of the peak current. Controllability is also affected by the switching frequency and it is recognized that a frequency ratio, $\alpha_f$, between the sampling and the synchronous frequency, lower than 10 might cause instability problems. As the synchronous frequency increases, the impact of the delay inherited by the sampling is getting more prominent resulting in deterioration of the reference frame of the current controller. Operating the EM at frequency ratios lower than 10 requires some kind of compensation for the time delays [46]. To be able to reduce the device losses and at the same time preserve system performance, a variable switching frequency can be used. A simple adaption of the switching frequency is introduced, where it is made proportional to the synchronous frequency of the EM and limited by a minimum frequency, $f_{sw,min}$, according to

$$f_{sw} = \max\left(f_{sw,min}, \alpha_f \cdot f_e\right)$$

(5.1)

where $f_e$ is the synchronous frequency of the EM. The frequency ratio, $\alpha_f$, is set to 10, according to the discussion above and $f_{sw,min}$ to 3 kHz, under the assumption that the torque degradation, due to increased ripple, is within acceptable limits.

5.5 Power Module Integration

Single-sided Cooling

The type of integration greatly affects the thermal impedance of the system, which can be seen in Figure 5.1, where the thermal impedances of module PM4 are shown for several different convection coefficients.
The thermal impedance for long pulse widths is highly dependent on the level of convection as well as the assembly. This dependency is gradually decreased for shorter pulse widths. Pulses shorter than approximately 50 ms, results in the same temperature rise, independent of the convection coefficient and assembly. The reason for this is that it takes a finite amount of time for the heat to propagate from the chip down to the bottom layer. The semiconductor chips is thereby becoming more and more thermally decoupled from the bottom module layers as the pulse width is decreased.

Figure 5.1 also shows that the combination of baseplate and heatsink results in a lower thermal impedance if the level of convection is low, i.e. below 5 kW/(m\(^2\)°C). The steady state thermal resistance is 0.40 °C/W when both baseplate and heatsink is used, whereas the thermal resistance is 0.44 °C/W and 0.41 °C/W in case of only using a baseplate or heatsink, respectively. The reason for this is that the combination of using a baseplate and heatsink increases the effective cooling area through a more effective heat spreading. Nevertheless, if the level of convection is high,
using both baseplate and heatsink, results in higher thermal impedance. The steady state thermal resistance for a convection coefficient of 50 kW/(m\(^2\)∙°C) is 0.33 °C/W when both baseplate and heatsink is used, whereas the thermal resistance is 0.26 °C/W and 0.20 °C/W in case of only using a baseplate or heatsink, respectively. As the level of convection is increased, the effect of heat spreading is gradually becoming less significant. Hence, the thermal impedance is lower for the assemblies with lower number of layers in the main heat path.

The choice of cooling directly on the baseplate or using a heatsink depends on the required level of convection. If the requirements are high, cooling directly on the baseplate is more beneficial in terms of lower thermal impedance.

**Double-sided Cooling**

The previous section describes how the thermal impedance is affected by the integration of single-sided cooled power modules. Increasing the convection coefficient reduces the thermal resistance. However, increasing the convection coefficient can be quite hard and expensive since it requires more advanced cooling techniques. The thermal resistance and its derivative for PM5 with respect the convection coefficient is shown in Figure 5.2(a) and (b), respectively.
Figure 5.2  Thermal resistance (a) of IGBT T1 for convection coefficients ranging from 0.5 to 50 kW/(m²°C), when equal amount of heat is dissipated in chip T11 and T12 of PM5b, PM5h and PM5hnb. Thermal resistance change with respect to convection coefficient (b).

The figure shows that the payoff in reduced thermal resistance is reduced significantly for more advanced cooling techniques. Hence, other solutions need to be considered. One way is to use materials with higher thermal conductivity. However, this option will not have any huge effect on thermal resistance. Using double-sided cooling on the other hand has, at least theoretically, the potential of lowering the thermal resistance with more than a factor of two. When referring to double-sided cooling, the solution with two substrates and no baseplate, described in Chapter Chapter 2, is assumed. The double-sided cooled version of PM5 has exactly the same chip layout as the single-sided cooled version. The only thing that differs is that the copper layers of the DBC:s are assumed to be thicker. The copper layers have a thickness of 250 μm in the original assembly, but is twice as thick, i.e. 500 μm, in the double-sided cooled assembly. The reason why the copper foil is made thicker, is that when the baseplate is omitted, some extra mechanical support for the ceramic
substrate is needed. The thermal resistance of PM5b and PM5dsc and the relative decrease in thermal resistance when using double-sided cooling is shown in Figure 5.3(a) and (b).

The figure shows that the thermal resistances, when using double-sided cooling, are lower for all convection coefficients. A relative decrease above 50% is only achieved for convection coefficients above 30 kW/(m²°C). The reason why it is not higher for the lower convection coefficients is that the heat spreading effect of the baseplate has a high impact. As has been mentioned before, it is better to compare the thermal impedances, since, compared to the thermal resistances, they reflect the thermal transient response of the module. The thermal impedance of PM5b and PM5dsc and the relative decrease in the thermal impedance when using double-sided cooling is shown in Figure 5.4(a) and (b).
Figure 5.4 Thermal impedance of IGBT T1 (a) for convection coefficients ranging from 1.0 to 30 kW/(m$^2$·°C), when equal amount of heat is dissipated in chip T11 and T12 of PM5b and PM5dsc. Relative thermal impedance change when using double-sided cooling (b).

Figure 5.4 (a) shows that the thermal impedance when using double-sided cooling is lower for all pulse lengths and convection coefficients above 5.0 kW/(m$^2$·°C). For the low convection coefficient, i.e. 1.0 kW/(m$^2$·°C), the thermal impedance is higher for pulse lengths ranging from 0.7 to 20s. The reason for this is again the reduced thermal inertia of the module due to the omitting of the baseplate.

### 5.6 Cooling Medium Temperature

The cooling medium temperature can be seen as an offset temperature for the power module. Increasing the cooling medium temperature increases the losses and reduces the available temperature margin in which the devices can be operated. The increased cooling medium temperature can, to some extent, be compensated by an increased convection coefficient. However, as was pointed out in the previous section, there is only a slight change in the thermal impedance for a large change in the convection...
Limited Regenerative Torque

Coefficient, and for pulses lower than 50 ms, there is hardly any change at all. The maximum cooling medium temperature as function of convection coefficient for different power module assemblies is presented in the results section, see Chapter 6.

5.7 Limited Regenerative Torque

One simple action to lower the cooling requirements is to limit the torque of the electrical machine during regenerative breaking. High breaking torques and rather low machine speeds results in high stress of the power module chips. Figure 5.5 shows an example of the diode junction temperature evolution for PM4h both with and without a limitation in the regenerative torque.

Figure 5.5 Temperature rise of chip D11 with and without a limited regenerative torque for a part of the FTP75 cycle.

The torque was limited from a minimum of -230 and -325 Nm to -150 Nm, for the two machines. This limitation results in an decrease of the cooling requirements, for the example presented in Figure 5.5, from 3.6 to 2.9 kW/(m²·°C), which can be considered as a rather modest change.
However, some of the highest temperature changes and peaks are avoided, which is advantageous from a reliability perspective. The drawback of limiting the regenerative torque, is that the amount of recovered energy is decreased. The FTP75 cycle is the cycle where the highest amount of energy will be lost due to the torque limitation. The recovered energy is decreased by 8% and 11%, equivalent to approximately 0.27 kWh and 0.41 kWh, for the two machines, respectively. This change in the recovered energy is assumed to be acceptable in order to lower the cooling requirements and enhance the reliability.
Chapter 6

Results

6.1 Maximum Junction Temperature

Power Module 1 and 5

In this section a comparison of the cooling requirements for two different inverters based on either three single phase modules or one three phase module is carried out. The inverters are assumed to be based upon PM1 and PM5, respectively. For details regarding materials, layout and dimensions, see Table 4.2. Figure 6.1 shows the required convection coefficients to keep the junction temperatures of the transistors below 125 °C for two different cooling medium temperatures, 60 °C and 80 °C. Both a fixed and a variable switching frequency, adapted to the synchronous frequency according to Equation (6), are used to derive the results presented in Figure 6.1(a-b) and (c-d), respectively.

The cooling level is, as has been stated before, considered to be sufficient if the junction temperatures of all the power module devices are kept below 125°C, in order to ensure long term reliability. With this constraint on the temperature, simulations show that if a fixed switching frequency is used, the inverter based on PM5 requires a convection coefficient of approximately 2.0 to 6.3 and 18 to 23 kW/(m²·°C), depending on the integration solution, for cooling medium temperatures of 60°C and 80°C, respectively. The inverter based on PM1 requires, in all cases, either higher or equal convection coefficients compared to the inverter based on PM5. The reason for this can be explained by studying the differences in the thermal impedances for the two modules. Figure 6.2 shows the thermal impedances for PM1 and PM5 for two different convection coefficients, 1 and 10 kW/(m²·°C) respectively.
Figure 6.1 Minimum convection coefficient, $h_{\text{conv, min}}$, to keep the junction temperature of the transistors below 125°C for two different cooling medium temperatures, 60 °C (a and c) and 80 °C (b and d), with fixed (a-b) and variable (c-d) switching frequency.
Figure 6.2 Thermal impedance of chip T11 for convection coefficients of 1.0 and 10 kW/(m$^2$·°C), when equal amount of heat is dissipated in chip T11 and T12 of PM1b and PM5b (a) together with PM1h and PM5h (b).
Chapter 6. Results

Three main things can be noted from the figures:

- The thermal impedance for PM1h is higher than PM5h for all convection coefficients both for the switch and the diode (see Figure 6.2b).
- The thermal impedance for PM1b is higher than PM5b for all convection coefficients for pulses shorter than 2s (see Figure 6.2a).
- The thermal impedance for PM1b is lower than PM5b for convection coefficients lower than 5 kW/(m²°C) and pulses longer than 2s (see Figure 6.2a).

The total thermal resistance computed from layer dimensions and material data is 0.19 °C/W and 0.27 °C/W for PM5b and PM1b, respectively. The thermal resistance is computed assuming no heat spreading, whereby the chip area has been used to compute all thermal resistances. The reason for the PM5b setup having a lower thermal resistance is that the PM5 module has a larger chip area than the PM1 module. The PM5 module also has a larger silicon volume, which results in a higher thermal capacitance and hence a lower thermal impedance. The higher thermal resistance and lower thermal capacitance of PM1 explains the two first points above. The third, and last, point is related to the geometrical layout of the modules. When $h_{\text{conv}}$ is low, below 5 kW/(m²°C), and baseplate cooling is employed, the geometrical layout of the module has a high impact on the thermal impedance. The layout of the PM1 module, see Appendix A, results in a lower thermal impedance since the chips are further apart, resulting in less thermal interaction between the chips. However, as $h_{\text{conv}}$ is increased and the heat spreading is becoming less and less prominent, the influence of the geometrical layout is also decreased and the actual thermal impedance of the chip itself is gradually becoming more important.

For the PM1 based inverter, the results show that if the cooling medium temperature is low (60°C), a heatsink should be used. On the other hand, when the cooling medium temperature is high (80 °C), direct baseplate cooling is more beneficial, since the cooling requirements are somewhat lower. When it comes to the PM5 module, using a heatsink is the preferred choice also for the higher cooling medium temperature of 80 °C. The
results from the simulations are summarized in Table 6.1 below.

<table>
<thead>
<tr>
<th>Power module nr 1 and 5</th>
<th>Fixed switching frequency</th>
<th>Variable switching frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transistor [kW/(m²·ºC)]</td>
<td>Diode [kW/(m²·ºC)]</td>
</tr>
<tr>
<td>T&lt;sub&gt;cl&lt;/sub&gt;</td>
<td>60 ºC</td>
<td>80 ºC</td>
</tr>
<tr>
<td>PM1b</td>
<td>5.4</td>
<td>18.1</td>
</tr>
<tr>
<td>PM1h</td>
<td>3.3</td>
<td>18.2</td>
</tr>
<tr>
<td>PM5b</td>
<td>6.3</td>
<td>22.7</td>
</tr>
<tr>
<td>PM5h</td>
<td>2.0</td>
<td>18.1</td>
</tr>
<tr>
<td>PM5hn</td>
<td>4.4</td>
<td>22.4</td>
</tr>
</tbody>
</table>

The results presented in Table 6.1 shows that all the requirements on the convection coefficient, when using a variable switching frequency, are lower. Reductions ranging from 20 to 80% are obtained, where the maximum decrease is obtained for PM5h and a cooling medium temperature of 80 ºC.

Increasing the cooling liquid temperature reduces the margin to the maximum temperature. Figure 6.3 shows the maximum cooling liquid temperature that can be used as a function of the convection coefficient for the FTP75 cycle. It should be pointed out that the data presented in this figure is derived from simulations with a cooling liquid temperature of either 60 or 80 ºC. Hence, no simulations for other than these temperatures are carried out so the results should be treated with some care. Since the losses increase with the temperature, the results presented in Figure 6.3 are somewhat optimistic. However, the results give an indication of the maximum cooling liquid temperatures that can be used for different levels of the convection coefficient and assemblies.
Figure 6.3 Maximum allowed cooling liquid temperature as a function of the convection coefficient, $h_{\text{conv}}$, for PM1 and PM5 using a fixed or variable switching frequency.

Figure 6.3 also shows that fewer number of layers within the module structure is beneficial as the convection coefficient goes up and a turning point can be seen around $20 \text{ kW/(m}^2\cdot\text{C})$. Using a variable switching frequency reduces the cooling requirements, and consequently a higher cooling liquid temperature can be tolerated. An increase of approximately 10-15 °C is possible, depending on the type of module and convection coefficient. In addition to this, the results show that having cooling liquid temperatures in the vicinity of 100 °C is not possible with these modules having a maximum admissible junction temperature of 125 °C. The margin for the temperature swings are simply too low.

**Power Module 3 and 4**

This section focuses on comparing the cooling requirements for a silicon (Si) and a silicon carbide (SiC) based power electronic inverter. The inverters are assumed to be based upon the PM3 and PM4 setups, see Table 4.2 for details. The same types of simulations as presented in the previous sections are carried out. Figure 6.4 shows the required convection
coefficients to keep the junction temperatures of the transistors below 125 °C for two different cooling medium temperatures, 60 °C and 80 °C. Both a fixed switching frequency and a variable, adapted to the synchronous frequency according to Equation 5.1, have been used to derive the results presented in Figure 6.4 (a-b) and (c-d), respectively.

Simulations show that if a fixed switching frequency is used, the Si based inverter requires a convection coefficient of approximately 2.5 to 4.5 and 9.5 to 16 kW/(m²·°C) depending on the integration solution, for cooling medium temperatures of 60°C and 80°C, respectively. The SiC based inverter requires much lower convection coefficients, 0.6 and 1.6 kW/(m²·°C). For the Si based inverter, the results show that if the cooling medium temperature is low (60°C), a heatsink should be used. On the other hand, when the cooling medium temperature is high (80 °C), the other two integration options are more beneficial, since the cooling requirements are much lower. The results clearly show that the SiC based inverter requires a much lower convection coefficient than the Si based inverter. The results from the simulations are summarized in Table 6.2 below.

<table>
<thead>
<tr>
<th>Power module nr 3 and 4</th>
<th>Fixed switching frequency</th>
<th>Variable switching frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Transistor [kW/(m²·°C)]</td>
<td>Diode [kW/(m²·°C)]</td>
</tr>
<tr>
<td>T_{cl}</td>
<td>60 °C 80 °C 60 °C 80 °C</td>
<td>60 °C 80 °C 60 °C 80 °C</td>
</tr>
<tr>
<td>PM4b</td>
<td>4.5 11.9 3.6 9.6</td>
<td>4.5 9.9 2.9 6.9</td>
</tr>
<tr>
<td>PM4h</td>
<td>2.5 15.7 1.7 13.3</td>
<td>2.5 14.6 1.3 5.5</td>
</tr>
<tr>
<td>PM4hnb</td>
<td>2.8 9.5 2.0 8.4</td>
<td>2.8 8.9 1.5 4.9</td>
</tr>
<tr>
<td>PM3h</td>
<td>0.6 1.6 0.8 2.4</td>
<td>0.6 1.6 0.7 2.0</td>
</tr>
</tbody>
</table>
Figure 6.4 Minimum convection coefficient, $h_{conv,min}$, to keep the junction temperature of the transistors below 125°C for two different cooling medium temperatures, 60 °C (a and c) and 80 °C (b and d), with fixed (a-b) and variable (c-d) switching frequency.
6.1. Maximum Junction Temperature

The results presented in Table 6.2 shows that all the requirements on the convection coefficient, when using a variable switching frequency, are either the same or lower. A maximum decrease of 60% is obtained for the Si based inverter with heatsink cooling and a cooling medium temperature of 80 °C.

As for the modules PM1 and PM5 it is interesting to study how much the cooling liquid temperature can be increased. Figure 6.5 shows the maximum cooling liquid temperature that can be used as a function of the convection coefficient for the FTP75 cycle.

![Figure 6.5 Maximum allowed cooling liquid temperature as a function of the convection coefficient, $h_{\text{conv}}$, for the three different assemblies, using a fixed or variable switching frequency.](image)

Once again the results show that less number of layers within the module structure is beneficial when a higher convection coefficient is used. The break point for the PM4 module is between 4 and 9 kW/(m²·°C), which is about half of the convection coefficient compared to the PM5 module.

The results presented in Figure 6.5 show that using a variable switching
frequency reduces the cooling requirements, and consequently a higher cooling liquid temperature can be tolerated. The results indicate that cooling liquid temperatures in the vicinity of 100°C could be used for the SiC based inverter, using a variable switching frequency.

**Double-sided Cooling**

This section focuses on comparing the cooling requirements for an inverter using single-sided and double sided cooling. The inverters are based on PM5. The same types of simulations as presented in the previous sections are carried out. Figure 6.8 shows the required convection coefficients to keep the junction temperatures of the transistors below 125 °C for two different cooling medium temperatures, 60 °C and 80 °C.

![Figure 6.6 Comparison of minimum convection coefficient, $h_{\text{conv},\text{min}}$, to keep the junction temperature of the transistors below 125°C for two different cooling medium temperatures, 60 °C (a) and 80 °C (b), with fixed switching frequency for PM5b and PM5dsc.](image)

The results show that using double sided cooling lower the cooling requirements significantly. The cooling requirements, for the most demanding cycle, when using double-sided cooling are approximately 1.4
6.1. Maximum Junction Temperature

and 2.5 kW/(m²·°C) for a cooling medium temperature of 60 and 80 °C, respectively. This is approximately 5 and 9 times lower, compared to using single-sided cooling.

Figure 6.7 shows the maximum cooling liquid temperatures that can be used without exceeding the maximum junction temperature of 125 °C. As was pointed out in the previous sections, comparing PM1 to PM5 and PM3 and PM4, the maximum cooling liquid temperatures, presented in Figure 6.7, should be treated with some care. The reason is that they are extrapolated from simulations where a cooling liquid temperature of 60 or 80 °C is assumed.

The results indicate that a significantly higher cooling liquid temperature can be tolerated when double-sided cooling is used. The difference ranges from 45 to 20 °C as the convection coefficient is increased. The results indicate that cooling liquid temperatures in the vicinity of 100 °C could be used for the double-sided cooled assembly.
Hill-hold

The sections above are dedicated to determine the cooling requirements for the modules based on several commonly used driving cycles. However, a feature that might determine the cooling requirements is hill-hold. This feature is here defined as standing still in a slope with the aid of any of the vehicle’s power sources. Hence, it is not assumed that the mechanical brakes are used. This feature is taken for granted for a conventional passenger car. However, for a plug-in hybrid, operating in pure electric mode, or a fully electric car, this feature implies tough thermal conditions for the power electronics. The required level of cooling that is needed for the hill-hold feature is determined by the time the vehicle should be able to stand still and the inclination of the slope. The required time is here assumed to be “infinite” and inclinations ranging from 4° to 20° are assumed. The reason that this feature causes high thermal stress on the inverter, and more specifically on the modules, is that only a limited number of the devices are exposed to the inverter losses. The current required to produce the torque that keeps the vehicle at stand still is actually a direct current, since the vehicle is not moving. This applies to vehicles with synchronous traction machines, whereas an induction type traction machine in the same situation would operate at slip frequency, typically a few Hertz. Depending on the absolute position of the EM rotor, it might happen that one phase is exposed to the full current and the other two phases are each carrying half of that current. The actual current level is determined by the required torque, which, in turn, is determined by the inclination. Since, the current is not changing sign, only the upper transistor and the lower diode of the phase leg are carrying the current. Hence, the total losses are only split between half of the total number of devices. If the vehicle is only slightly moving, or an induction type of traction machine is used, the losses are spread between all devices, and the maximum junction temperature is decreased. This can clearly be seen in Figure 6.8 where the results from two simulations of a hill-hold are presented.
6.1. Maximum Junction Temperature

In the first simulation the vehicle is standing completely still, whereas in the second simulation it is moving slowly forward after 30s at a speed of 20 cm/s. As the losses are distributed more evenly within the module when the vehicle is moving, the temperature drops approximately 15° within 5s.

Several simulations with different inclination angles are carried out in order to study how it affects the cooling requirements and at what level the cooling requirements are actually determined by the hill-hold feature. The results for several simulations with different inclination angles are presented in Figure 6.9.
Figure 6.9 Maximum IGBT junction temperatures for PM5b and PM5h during hill-hold for different inclination angles.

The figure shows that the inclination angle has a significant impact on the required convection coefficient. As an example it can be mentioned that the minimum convection coefficient for PM5h has to be increased by a factor of 10, from 2.0 to 20 kW/(m²·°C), if the inclination angle is changed from 12° to 20°. One way to lower the junction temperatures, which is discussed extensively in previous chapters, is to use a variable switching frequency. Figure 6.10 shows the minimum required convection coefficients for PM5 for several hill-hold cycles with different inclination angles, using both a fixed and variable switching frequency.
6.1. Maximum Junction Temperature

Figure 6.10 Minimum required convection coefficient to keep the IGBT junction temperatures for PM5 below 125 °C, with a fixed (a) and variable (b) switching frequency.

Using a variable switching frequency lowers the cooling requirements substantially. For an inclination angle of 20°, depending on the integration of the module, the required convection coefficient is lowered with a factor of 3 to 4. The results from the hill-hold simulations show that the required cooling is highly dependent on the inclination angle requirements. Comparing the results presented in Figure 6.10 and Table 6.1, shows that the hill-hold feature is actually determining the cooling level if an infinite hill-hold time at inclination angles above 12° is required, assuming a constant switching frequency. Using a variable switching frequency lowers the requirements on the cooling and inclinations up to 16° can be handled.

**Loss Sensitivity**

The power electronic losses used in the simulations are computed from datasheet values according to the procedure presented in Chapter 4.5. Computing the losses in this way basically assumes that the external conditions such as the DC-link voltage and the gate-emitter voltage are the same as in the datasheet. It also assumes that the hardware is the same, i.e.
that it has the same stray inductance and gate resistor. Since this, most probably, is not the case, a sensitivity study regarding the losses are carried out. The aim of this study is to examine how much influence the loss modelling accuracy has on the final cooling requirement. Simulations were the losses have been both decreased and increased has been carried out. It is assumed that the highest uncertainty in the loss modelling can be found in the switching losses, due to its dependency on stray inductance, gate resistor, DC-voltage and reverse recovery current of the diode. The loss change is hence distributed in a 5\% change in the conduction losses and 15\% in the switching losses. Figure 6.11 show the cooling requirements for the different assemblies of PM4 and PM1 for three different levels of the losses.

The simulations show that there is no significant change in the cooling requirements as the losses are changed. The maximum junction temperatures for the switches are found in the end of periods of heavy accelerations. As the speed increases during an accelerations the electrical frequency increases and with that the voltage applied to the EM. Higher voltages results in an increased duty cycle for the switch and since the conduction losses are proportional to the duty cycle, they are also increased. Thus, the major part of the total losses is conduction losses and the 15\% increase in the switching losses does not have any significant impact on the maximum junction temperature.
Figure 6.11 Minimum convection coefficient, $h_{\text{conv, min}}$, to keep the junction temperature of the transistors and diodes for PM4 (a-b) and PM1 (c-d) below 125°C for three levels of the losses, respectively.


6.2 Reliability Results

As was stated in the beginning of this chapter, deciding the cooling requirements only based on not exceeding an admissible maximum junction temperature will most probably have a negative impact on the power module reliability. The reason for this is that no concern is taken to the thermal cycling of the device. The results presented in this chapter are based on the simulations presented in Chapter 6, and extended with a reliability analysis. The first part of the analysis studies how the reliability, in terms of the number of cycles until failure, is influenced by the convection coefficient. The second part discusses how the use of a variable switching frequency can extend the life time of power modules. The third and last part compares the differences in the reliability between the one and three phase modules.

Reliability Estimation

As was described in the chapter regarding the system simulations, several simulations for different driving cycles and convection coefficients have been carried out. The results from these simulations are used as input to the reliability models presented in Chapter 2.2, to estimate the reliability of the different power modules. Figure 6.12(a) shows the results of the reliability analysis for PM5b in terms of the number of missions until failure as a function of the convection coefficient. In addition to this, Figure 6.12(b) shows how many times the lifetime is increased, for each cycle, when the convection coefficient is changed from the lowest possible value.

The results clearly show that much can be gained, in terms of reliability, by increasing the convection coefficient. Taking the US06 cycle as an example, a convection coefficient of approximately 3 kW/(m²·ºC) is required to keep the junction temperature below 125 ºC. However, increasing the convection coefficient with a factor of 3, i.e. from 3 to 10 kW/(m²·ºC), results in an increase of the number of missions until failure with a factor of 8 and 4 for the LESIT and CIPS08 model, respectively. Increasing the cooling further, is increasing the lifetime, however there is a significant drop in the relative gain increasing it above 30 kW/(m²·ºC).
6.2. Reliability Results

Figure 6.12 Number of missions until failure for PM5b as a function of convection coefficient and driving cycle (a). The number of times the lifetime is increased by changing the coefficient from the lowest possible value (b).

If the number of missions until failure estimated by the LESIT and CIPS08 model is compared, two distinctions are seen. The LESIT model estimates a higher number of missions until failure for the less demanding driving cycles, such as the NEDC. A high convection coefficient imposes better cooling and thereby lower mean temperatures and temperature swing. Hence, for high convection coefficients the LESIT model estimates a higher number missions until failure. This is clearly seen in Figure 6.1, where the number of missions until failure, estimated by the LESIT model, is lower for a convection coefficient of 5-30 kW/(m²°C) and higher above 30 kW/(m²°C). This refers back to what is stated in Chapter 2.2, that the LESIT model estimates higher lifetimes for a range of combinations where the temperature swings and mean temperatures are low and this range is growing as the on-time is increased.

The significant increase in lifetime as the cooling is increased is a consequence of two things, lower temperature swings at lower mean
junction temperatures. Figure 6.13(a) and (c) shows the frequency of the different temperature swings and their associated mean temperatures for the FTP75 cycle. Note that temperature swings lower than 5 °C are removed. The reason for that is that the majority of the temperature swings are below 5 °C and the swings above would not be seen in the bar graphs.

The figure clearly shows how the temperature distribution is shifted towards the left corner, i.e. towards lower temperature swings and mean temperatures, as the convection coefficient is increased. Figure 6.13(b) shows the relative distributions of the damage caused by the temperature combinations presented in Figure 6.13(a). The relative distributions, $R_{lc,ij}$,
6.2. Reliability Results

are derived by using Equation 2.1 and 2.4 according to

\[
R_{lc,ij} = \frac{\left( c_{lc} \right)_{i,j}}{c_{lc,\text{tot}}} = \frac{\left( n_{lc} \right)_{i,j}}{\sum_{i,j} \left( n_{lc} \right)_{i,j}} \tag{6.1}
\]

where \((c_{lc})_{i,j}\) is the damage caused by the \(i\)th mean temperature and the \(j\)th temperature excursion for a low convection coefficient. The indexes \(lc\) and \(hc\) indicate a low and high convection coefficient of 5 and 50 kW/(m²·°C), respectively. The parameter \(c_{lc,\text{tot}}\) is the accumulated damage for all of the cycles presented in Figure 6.13. The parameters \(\left( n_{lc} \right)_{i,j}\) and \(\left( N_{f,lc} \right)_{i,j}\) is the number of cycles and the number of cycles until failure for the \(i\)th mean temperature and the \(j\)th temperature excursion. The relative distributions for low convection coefficients, \(R_{lc,ij}\), are then finally expressed in percent. The relative distributions for high convections coefficients are computed in the same way, but the individual damage, \((c_{hc})_{i,j}\), is scaled with the total accumulated damage for the case with low convection coefficient according to

\[
R_{hc,ij} = \frac{\left( c_{hc} \right)_{i,j}}{c_{lc,\text{tot}}} \tag{6.2}
\]

The reason for using \(c_{lc,\text{tot}}\) is to show how much smaller the actual damage is. The total accumulated damage for the case with high convection coefficient is only around 2% of that with a low convection coefficient. Figure 6.13(b) shows that the almost 50% of the total damage caused by the presented cycles are due to the 4 cycles occurring at \(T_m=86\) °C and \(\Delta T=70\) °C. The results above show the importance of keeping the mean temperatures and the temperature excursions as low as possible in order to achieve a long module lifetime.

In Chapter Chapter 6 it is shown that the cooling requirements can be significantly decreased by using a variable switching frequency, due to the lower switching losses. Hence, it is reasonable to think that keeping the cooling level and changing from a fixed switching frequency of 10 kHz to
Chapter 6. Results

A variable switching frequency increases the number of missions to failure. Figure 6.14 shows that this is the case, presenting the number of times the lifetime is increased for each driving cycle and level of cooling.

Figure 6.14 Number of times the lifetime is increased by changing from a fixed to variable switching frequency for PM5b.

Figure 6.14 shows that the relative gain in lifetime, when using a variable switching frequency, is different when comparing the different cycles. For the PM5b module a relative gain of approximately 8 to 10 and 4 to 6 can be achieved for the HYZEM Rural cycle, using the LESIT and CIPS08 model, respectively. The result for the HYZEM Highway cycle indicate a much lower increase in the lifetime, where the gain is only between 2 and 3 for both of the reliability models. This difference in the relative gain in lifetime is due to the different characters of the cycles. The HYZEM Highway cycle is a cycle with rather high constant speed, which implies little gain in switching losses since the switching frequency is adapted to the vehicle speed. The HYZEM Rural cycle on the other hand has portions of low speed and the losses are thereby lowered significantly. The reason for this can be explained by studying Figure 6.15.
Figure 6.15 IGBT junction temperature and the associated losses for the HYZEM Rural cycle, for a fixed (a-b) and variable switching frequency (c-d) respectively. IGBT junction temperature and the associated losses for the HYZEM Highway cycle, for a fixed (e-f) and variable switching frequency (g-h) respectively.
Figure 6.15(a-b) and (c-d) shows the IGBT junction temperature and the associated losses for the HYZEM Rural cycle, for a fixed and variable switching frequency, respectively. The same thing is shown in Figure 6.15(e-f) and (g-h) for the HYZEM Highway cycle. Just by studying Figure 6.15 it can be seen that the relative change in the losses, when switching to a variable switching frequency, is significantly more evident for the HYZEM Rural cycle. The higher relative decrease in the losses results in higher relative decrease in the temperature swings, which in the end leads to a higher relative gain in the lifetime of the device.

It was shown in Chapter 5.5 that the module assembly greatly influences the thermal impedances of the power module. Hence, the reliability, in terms of the number of cycles until failure, should also be significantly affected. Figure 6.16 shows the number of missions until failure for the NEDC and FTP75 cycles for the different assembly options stated in Table 4.1 using a fixed (a) and variable (b) switching frequency.

The differences in the number of missions until failure are a direct consequence of the differences in the thermal impedance for the different module assemblies. The figure shows that the heatsink option is the most beneficial when the convection coefficient is below 20-30 kW/(m²·°C). If a higher convection coefficient is used, cooling directly on the baseplate is the best option. This is totally in line with the discussion regarding the thermal impedance and the turning point, i.e. the convection coefficient level where the heatsink option no longer is the most beneficial, is about the same.
6.2. Reliability Results

Figure 6.16 Number of missions until failure for the NEDC and FTP75 cycles for the different cooling options stated in Table 4.1, using a fixed (a) and variable (b) switching frequency.
In Chapter 6 the cooling requirements for keeping the junction temperature below 125 °C for inverters based on PM1 and PM5 are compared. Here, this comparison is extended with an analysis of the reliability of the two different modules. Figure 6.17 shows the number of missions until failure as a function of the convective coefficient.

Figure 6.17 Number of missions until failure comparison of PM1b/PM5b (a) and PM1h/PM5h (b).

Figure 6.17(a) shows that PM1b can sustain a higher number of missions compared to PM5b. This is possible despite the fact that PM1b has higher thermal impedance than PM5b, which was shown in Chapter 4.6, since PM1 has lower losses. Hence, the somewhat higher thermal impedance of PM1b is overridden by the lower losses. The opposite can be seen in Figure 6.17(b), where the lower thermal impedance of PM5h results in a more beneficial thermal evolution, from a reliability point of view, even though the losses are higher. The absolute difference in thermal impedance between PM1b and PM5b is in the same range as the difference for PM1h and PM5h. However, the relative difference is higher when using a heatsink, since the actual thermal impedance is lower.
Chapter 7

Concluding Remarks and Future Work

The aim of this thesis is to determine if it is possible to integrate the power electronics and the electric machine into one unit. To be able to give an answer to that, several electro-thermal simulations are carried out in order to determine the cooling requirements of the inverter. The foundation for these simulations is the simulation model of the EDS, where the main focus is on the thermal model of the inverter itself. Using a FEA-tool to develop simplified thermal models, in terms of impedance matrices, is proven to be a very effective way of simulating the thermal behaviour of power modules. A systematic way of determining the cooling requirements with the aid of the developed thermal models is proposed.

7.1 Summary of Findings

Thermal modelling of power modules in combination with electro-thermal simulations is the main focus of this thesis. Some main conclusions regarding the thermal modelling are drawn.

First, the importance of using a thermal model that takes the thermal coupling between single chips in a power module into account is pointed out. Neglecting the thermal coupling between the chips can, for low convection coefficients, give rise to significant underestimation of the junction temperatures. As the convection coefficient is increased, the thermal coupling is becoming less pronounced, the model error is reduced. For the modules studied in this thesis, an acceptable level of the model error is obtained if the convection coefficient is kept above 30 kW/(m²·°C).

As the convection coefficient is increased the thermal resistance is
lowered. The thermal impedance does also change but only for pulse widths greater than approximately 0.05 to 1 s depending on the module assembly. The type of module assembly greatly influences the thermal impedance. The developed thermal models show that as the convection coefficient is low, i.e. lower than approximately 5 kW/(m²·ºC), using a heatsink results in the lowest thermal resistance. As the convection coefficient is increased omitting the heatsink or omitting the baseplate and mount the substrate directly on the heatsink results in the lowest thermal resistance. This is basically a trade-off between heat spreading and a high thermal resistance in the main heat path. However, omitting either the heatsink or the baseplate reduces the thermal inertia and results in a higher thermal impedance for some pulse lengths. Hence, determining what assembly to use, only by studying the thermal impedance, might not result in the best choice. It is the combination of the assembly and the usage of the power module that needs to be considered and matched.

Increasing the convection coefficient is associated with using more advanced cooling techniques, which often results in a more expensive solution. In addition to this, it is shown that the payback in terms of a reduction in the thermal resistance is significantly reduced as the convection coefficient reach levels above 3 kW/(m²·ºC). However, using the double-sided cooled assembly described in Chapter Chapter 2 is shown to have significant impact on the thermal impedance even for convection coefficients above 3 kW/(m²·ºC). As a matter of fact, it is above this level, where the double-sided cooled assembly has its highest potential. It is shown that it is actually able to reduce, not only the thermal resistance, but also the thermal impedance with up to 50%.

The thermal models are used in a simulation model of the complete EDS, together with a rather simple vehicle model in order to study the thermal cycling of the power modules. The criterion for a sufficiently high convection coefficient is that the junction temperature of all the devices within the power module, should not exceed 125 ºC.

For the 400A power modules the results show that a convection coefficient in the range of 2 to 6 kW/(m²·ºC) is required if the cooling medium temperature is assumed to be 60 ºC. Allowing a higher cooling medium temperature of 80 ºC requires a higher convection coefficient, in the range of 18 to 23 kW/(m²·ºC). Using a variable switching frequency can reduce
the cooling requirements significantly. A reduction of 20 to 80% is shown to be possible, depending on the module assembly and cooling medium temperature. The maximum cooling medium temperature as function of the convection coefficient is presented and the results indicate that the cooling medium temperature should be kept below 80 and 90 °C, if a fixed and variable switching frequency is used, respectively.

For the 300A silicon based power module the results show that a convection coefficient in the range of 2.5 to 4.5 kW/(m²·°C) is required if the cooling medium temperature is assumed to be 60 °C. The same figure for the SiC based module is 0.8 kW/(m²·°C). The cooling requirements for the SiC module is much lower and this is even more evident if a higher cooling medium temperature of 80 °C is assumed. The required convection coefficients are then 8.4 to 13.3 kW/(m²·°C) and 2.4 kW/(m²·°C) for the Si and SiC based module, respectively. Using a variable switching frequency reduces the cooling requirements for the Si based module with up to 60%. Whereas for the SiC based modules, there is hardly any change at all due to the inherently low switching losses. The maximum cooling medium temperature as function of the convection coefficient is presented and the results indicate that the cooling medium temperature should be kept below 85 and 95 °C for the Si based module, if a fixed and variable switching frequency is used, respectively. The same figures for the SiC based inverter are 97 and 107 °C.

It is stated that using double-sided cooling has significant impact on the thermal impedance, which can be reduced with up to 50%. The consequence of this reduction is clearly seen as a significant reduction in the cooling requirements. For a cooling medium temperature of 60 °C, the required convection coefficient is 5 times lower, 1.4 compared to 6.3, when double-sided cooling is compared to single-sided baseplate cooling. The reduction is even higher for a cooling medium temperature of 80 °C, where the reduction is as high as 9 times.

Hill-hold implies tough thermal conditions for the power electronics and causes high thermal stress on the inverter, and more specifically on the modules. The results show that if a hill-hold feature with an “infinite” stall time is required, it might be determining the cooling requirements. If inclination angles above 12º are required, the required convection coefficient is determined by the hill-hold feature if a fixed switching
frequency is used. If a variable switching frequency is used inclination angles up to 16º can be handled.

In addition to determine the cooling requirements based on keeping all junction temperatures within the model below 125 ºC, a reliability analysis is carried out. The reliability analysis is based on the two most commonly used reliability models, namely the LESIT and the CIPS08 model. The result show that much can be gained in terms of an extended module lifetime, by increasing the convection coefficient above the level that is required to keep the junction temperature below 125 ºC. As an example, increasing the convection coefficient from 3 to 10 kW/(m²·ºC) for PM5 increases the lifetime with a factor of 4 to 8, depending on which reliability model that is used and which cycle that is studied.

The results clearly show that the lifetime can be increased by using a variable switching frequency. However, the level of the increase is dependent on the character of the driving cycle. For cycles with high constant speed, the benefit is not as high as for the cycles where the speed is lower and fluctuating. For PM5b the increase in lifetime, when using a variable switching frequency, is approximately 2 to 3 for the HYZEM Highway cycle, whereas the increase is 8 to 10 or 4 to 6 for the HYZEM Rural cycle, depending on the reliability model.

It is shown that the type of module assembly greatly influences the thermal impedance, and consequently the reliability of the module. For module PM5 it shown that an assembly with a heatsink is the preferred choice if the convection coefficient is below 20 to 30 kW/(m²·ºC).

A comparison of PM1b and PM5b shows that the former can sustain a higher number of cycles until failure, although having higher thermal impedance. The reason is simply that PM1b has lower losses, and the difference in losses is higher than the difference in the thermal impedance. However, comparing PM1h and PM5h shows that later can sustain a higher number of cycles until failure, which is the opposite to the comparison of PM1b and PM5b. The reason for this is that the significantly lower thermal impedance of PM5h outperforms the lower losses of PM1h.

The first objective mentioned in the thesis, is if it is possible, from a
thermal perspective, to integrate or attach the inverter to the electric machine. The answer to this question is in the results presented above. Yes, it must be considered possible, since the results do not point in the direction that any extraordinary capacities for the cooling system are required. However, it must be pointed out that the focus has been on the thermal environment for the power modules, nothing is said about the required installation space and the passive components.

7.2 Future Work

This thesis focuses on thermal modelling of power modules in a hybrid vehicle application. A hybrid electric vehicle has, apart from the internal combustion engine, at least one additional source of power available for vehicle propulsion. The control of these power sources is managed by an energy management controller and its associated energy control strategy. The control strategy has several different objectives, where the primary ones usually aim to minimize the fuel consumption and emissions. The control objectives are achieved by instantaneous management of the power split between the power sources. It is the author’s belief that this power split can be done in a better way if limitations in the EDS can be predicted beforehand. Such limitations could be, for example derating of the maximum output power of the inverter due to thermal limitations. Hence, knowledge about the actual junction temperature and being able to estimate the junction temperature evolution for some time horizon could be valuable in order to improve the energy control strategy. Power modules are normally equipped with a temperature sensor. However, these sensors are normally mounted on the substrate or the baseplate, far away from the chips. The actual temperature that is measured will therefore be significantly different from the actual junction temperatures. The proposed solution to this problem is to develop an observer based on a thermal model of the power module. The observer would then use the sensor information together with the measured currents as inputs. However, if this information is sufficient and if it is possible to develop a simple but accurate enough thermal model to run in a real-time control system and accurately estimate the junction temperatures, is left for future research.

Since the actual junction temperatures are normally not known, derating of the inverter power output is a difficult task. Knowing that the sensor is not measuring the junction temperature might result in a too restrictive
derating strategy, and an under usage of the inverter capacity. It might also result in the opposite, that the junction temperature frequently exceeds its upper limit, resulting in a decreased module lifetime. In lifetime testing of power modules temperature dependent parameters, such as the on-state voltage, are used as an indirect method of measuring the chip temperature. Including measurement electronics into a gate driver, in order to measure the on-state voltage would be of great interest. This would, with some signal processing, give an instant value of the mean value of the junction temperature of the chip. This information would be a valuable input to the observer mentioned above. It could also be used to detect the end of life of the module.

One limitation of the thermal models presented in this thesis is that they do not take into account any variation of material parameters, such as for example thermal conductivity, due to the change in temperature. This simplification does not introduce any large errors if the module is based on silicon, since the thermal conductivity has a rather low temperature dependency and the actual temperature range is rather low, around 100°C. However, for silicon carbide based devices, this might be an issue, since the temperature dependence of the thermal conductivity is stronger and the devices can be operated in wider temperature range.

An interesting direction for future work, however a bit off topic, would be to apply the method of the thermal modelling presented in this thesis on an electrical machine. This would be challenging since electric machines have more complex geometries and more sources of power loss. To complicate it further, some of the power loss sources are both temperature and speed dependent. In addition to this, the convection coefficient in the air-gap is speed dependent, which might need to be taken into account.
References


References


109

References


Appendix A

PM1 - Three Phase 400 A IGBT-module

Figure A.1 Self and mutual impedances for a convection coefficient of 1,0 kW/(m²∙°C), when equal amount of heat is dissipated in chip T11 and T12 of PM1b (a) and PM1h (b).
112

Appendix A

Figure A.2 Thermal impedance of chip T11 for convection coefficients ranging from 1.0 to 10 kW/(m^2°C), when equal amount of heat is dissipated in chip T11 and T12 of PM1b and PM1h.

PM3 - Single Phase 300 A SiC BJT module

Figure A.3 Self and mutual impedances for a convection coefficient of 1.0 kW/(m^2°C), when equal amount of heat is dissipated in chip T11 to T16 of PM3h.
Figure A.4 Thermal impedance of chip T11 for convection coefficients ranging from 1.0 to 10 kW/(m²·°C), when equal amount of heat is dissipated in chip T11 to T16 of PM3h.

PM4 - Single Phase 300 A IGBT-module

Figure A.5 Self and mutual impedances for a convection coefficient of 1.0 kW/(m²·°C), when equal amount of heat is dissipated in chip T11 and T12 of PM4b (a), PM4h (b) and PM4hnb (c).
Figure A.6 Thermal impedance of chip T11 for convection coefficients ranging from 1.0 to 10 kW/(m$^2$·°C), when equal amount of heat is dissipated in chip T11 and T12 of PM4b, PM4h and PM4hnb.

**PM5 - Single Phase 400 A IGBT-module**

Figure A.7 Self and mutual impedances for a convection coefficient of 1.0 kW/(m$^2$·°C), when equal amount of heat is dissipated in chip T11 and T12 of PM5b (a), PM5h (b) and PM5hnb (c).
Figure A.8 Thermal impedance of chip T11 for convection coefficients ranging from 1.0 to 10 kW/(m²·°C), when equal amount of heat is dissipated in chip T11 and T12 of PM5b, PM5h and PM5hnb.
Appendix B

\[ A_y = \begin{bmatrix}
-\frac{1}{\tau_{ij}} & 0 & 0 & 0 & 0 \\
0 & -\frac{1}{\tau_{ij}} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\frac{1}{\tau_{ij}} & 0 \\
0 & 0 & 0 & 0 & -\frac{1}{\tau_{ij}} \\
\end{bmatrix}, \quad (N_x \times N_y) \]  \hspace{1cm} (B.1)

\[ B_y = \begin{bmatrix}
R_{ij}^1 \\
\vdots \\
R_{ij}^{N_y} \\
\end{bmatrix}, \quad (N_x \times 1) \]  \hspace{1cm} (B.2)

\[ C_y = [1 \quad \cdots \quad 1], \quad (1 \times N_y) \]  \hspace{1cm} (B.3)

\[ A_y = \begin{bmatrix}
A_{ij} & 0 & 0 & 0 & 0 \\
0 & A_{ij} & 0 & 0 & 0 \\
0 & 0 & A_{ij} & 0 & 0 \\
0 & 0 & 0 & A_{ij} & 0 \\
0 & 0 & 0 & 0 & A_{ij} \\
\end{bmatrix}, \quad \left( \sum_{i=1}^{N_y} N_x(i) \times \sum_{i=1}^{N_y} N_x(i) \right) \]  \hspace{1cm} (B.4)
\( B_j = \begin{bmatrix} B_{ij} & 0 & 0 & 0 & 0 \\ 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & B_{ij} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & B_{ij} \end{bmatrix}, \left( \sum_{i=1}^{N} N_k(i) \times N_j \right) \) \hspace{1cm} (B.5)

\( C_j = [C_{ij} \quad \cdots \quad C_{ij}], \quad \left( 1 \times \sum_{i=1}^{N} N_k(i) \right) \) \hspace{1cm} (B.6)
Appendix C

VEHICLE PARAMETERS

Table C.1 Vehicle parameters.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( m_v ) [kg]</th>
<th>( C_d )</th>
<th>( C_r )</th>
<th>( \rho_v ) [kg/m(^3)]</th>
<th>( A_v ) [m(^2)]</th>
<th>( r_w ) [m]</th>
<th>( g ) [N/kg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1770</td>
<td>0.26</td>
<td>0.0118</td>
<td>1.225</td>
<td>2.16</td>
<td>0.3351</td>
<td>9.82</td>
</tr>
</tbody>
</table>

RELIABILITY MODEL PARAMETERS

Table C.2 Reliability model parameters.

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>( A ) [K^-(\alpha)]</th>
<th>( \alpha )</th>
<th>( E_a ) [J]</th>
<th>( k_B )</th>
<th>( \beta_1 ) - ( \beta_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>LESIT</td>
<td></td>
<td>302500</td>
<td>-5.039</td>
<td>9.891·10(^{-20})</td>
<td>1.3807·10(^{-24})</td>
<td>-</td>
</tr>
<tr>
<td>CIPS 08</td>
<td></td>
<td>9.30·10(^{14})</td>
<td>10</td>
<td>6</td>
<td>300</td>
<td>-4,416 1285</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>( K ) [A]</th>
<th>( I ) [V/100]</th>
<th>( V ) [V/100]</th>
<th>( D )</th>
<th>( \beta_1 ) - ( \beta_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-4,416 1285</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0,463 0,716</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0,761 -0,5</td>
</tr>
</tbody>
</table>

119
Appendix D

Figure D.1 Power module layout for PM1.

Figure D.2 Power module layout for PM3.
Figure D.3  Power module layout for PM4.

Figure D.4  Power module layout for PM5.