© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Digital Object Identifier (DOI): 10.1109/EPE.2014.6910822

Power Electronics and Applications (EPE’14-ECCE Europe), 2014 16th European Conference on

Review of active thermal and lifetime control techniques for power electronic modules

Markus Andresen
Marco Liserre
Giampaolo Buticchi

Suggested Citation
Review of Active Thermal and Lifetime Control Techniques for Power Electronic Modules

Markus Andrensen, Marco Liserre, Giampaolo Buticchi
Chair of Power Electronics
Christian-Albrechts Universität zu Kiel
Kaiserstraße 2,
Kiel, Germany
Tel.: +49431 /880 –6107
Fax: +49431 /880 –6103
E-Mail: ma@tf.uni-kiel.de, ml@tf.uni-kiel.de, gibu@tf.uni-kiel.de
URL: www.pe.tf.uni-kiel.de

Acknowledgements
The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 616344 – Heart.

Keywords
«Reliability 1», «Active thermal control 2», «Power cycles 3», «Thermal cycles 4».

Abstract
Lifetime of power electronics modules can be extended with passive methods (condition monitoring) and active ones. This paper intends to give an overview in the second category, namely active thermal control or lifetime control, offering a critical comparison based on a comprehensive reference list. Mission profiles are compared to evaluate the potential of the controllers.

I. Introduction

Some industrial applications like electrical vehicles or aircrafts are subject to very high safety regulations, which set high demands to the reliability of the electrical system [1]. For this reason the failures need to be minimized, whereby the power electronic modules offer potential improvements [2]. During operation, the power modules undergo power cycles and ambient temperature changes, which create junction temperature fluctuations in the module. These temperature fluctuations cause mechanical stress and they are the main reason for aging and failures. Thereby the stress depends on the maximum temperature, the average temperature and the magnitude of the thermal cycles. With the development of power electronic semiconductors made of wide bandgap materials such as 4H-silicon carbide (4H-SiC) and Gallium Nitride (GaN), it is possible to archive higher operating temperatures compared to Silicon (Si) based modules [3]. Thus even higher thermal cycling magnitudes are possible and a temperature management gets more important. Furthermore the packaging size of the Si-modules is decreasing, which makes the heat dissipation more challenging [4]. To ensure a reasonable module lifetime, the temperature has to be kept as low as possible and thermal cycles have to be either reduced in magnitude or in the total number of cycles.

In wire-bonded power modules with direct-bond-copper (DBC), the most relevant in the actual scenario, the main failures are bond wire lift off and solder fatigue [5]. Several possibilities exist to either replace the bond wires in new module structures, to control the module cooling [6], or to reduce the thermal cycling with software [7],[8] There are also materials with the capability to withstand extreme temperature cycles [9], but for most applications, this is not a cost efficient solution.

This work gives an overview on lifetime extension for power electronic modules, whereby aspects from hardware and software design are considered. First, the power module failures are described and a brief introduction in lifetime models is given in section II. Next, a review on various concepts for lifetime extension is given in section III and approaches for active thermal control are compared based on four typical mission profiles. Finally, in section IV challenges for active and passive thermal control are described.
II. Failures and lifetime estimation for power modules

Power electronic modules in the medium and high power range usually consist of Insulated Gate Bipolar Transistors (IGBTs) and Metal-Oxide Semiconductor Field Effect Transistors (MOSFETs) [10]. In the modules, the chips are mounted on multi layer configurations (DBC substrate) to ensure the electrical insulation and good heat dissipation [11]. Thereby the chip is soldered on the DBC, which consists of a substrate enclosed by two separated layers of copper. The complete structure for a module with baseplate is shown in Fig. 1. Normally the ceramic substrate consists of Al₂O₃ or AlN. Beside its lower heat transfer capability compared to copper, these materials have unequal heat expansion coefficients [5]. The resulting problem is the strain between the layers caused by temperature changes. This strain is regarded as the main reason for aging of power electronic modules in literature [1]. Thus temperature changes of the components need to be prevented for reduced strain and increased lifetime. The typical structure of a module is shown in Fig. 1.

An often addressed problem is the interconnection between the semiconductors. The normally used bond wires lift off in case of a soldered connection [1]. This is being improved with sintered connections, but heel cracking remains a common failure [12]. Another problem is solder fatigue, which either occurs between the chip and the DBC or between the DBC and the baseplate. This reduces the heat transfer capability, which results in higher junction temperatures and finally in a failure [13]. For protection against corrosion and environmental influences, a gel filling immerses the module to guarantee dielectric strength [14]. The main aging processes of this silicone gel are water trees, partial discharge and electrical trees, which are degrading effects causing aging but not the destruction [15].

Module manufacturers provide estimations for the lifetime of their devices in dependence of the number of thermal cycles the modules have to undergo during their lifetime. In Fig. 2 is presented a typical diagram showing the industrial standard and the lifetime extension of improved interconnections. It can be seen that the number of cycles \( N_f \) depends exponentially on the magnitude of the thermal cycles \( \Delta T \). This dependence is also shown in many lifetime models presented in literature, such as the Coffin Manson equation (1). The coefficients \( a \) and \( n \) are constants, which are empirically obtained.

\[
N_f = a \cdot (\Delta T)^{-n}
\]  

(1)

Since this model is regarded as too simple, there are several improvements, which add parameters like the average junction temperature, the maximum junction temperature, the frequency of the temperature cycles and further power module characteristics [11]. Nevertheless general use of these models needs to be regarded carefully [17],[18]. Other approaches are stress based, plastic strain based, creep strain based, energy based or damage accumulation based [19].

III. Approaches for lifetime extension

In general, the physics of failure approach is gaining importance in reliability research, because it provides a clear understanding of the potential improvement, which is not given by statistic values [20]. Examples for this approach are a detailed stress analysis of the thermal cycles and the calculation of the according stress and the consumed lifetime [21],[22].
A) System design

For lifetime extension of the devices, the temperature cycles have to be reduced in total number or in amplitude. An overview on possible improvements in the hardware design is given in Fig. 3. For the manufacturer, there is a high potential to improve the module packaging. A promising technology is 3D packaging, which would enable double sided cooling. This would improve the heat dissipation and thus reduce the maximum temperature compared to one sided cooling. Another potential for the manufacturers is to improve the interconnection between the chips and thus to reduce the chances of bondwire liftoff. Present technologies, which have been proven for lifetime extension, are heavy copper wire bonding, sintered connections or buffer layers between the materials [16][23][24].

Supplementary, the choice of the inverter topology offers potential, because multilevel topologies facilitate to share the stress between the devices. Thus the stress of a single chip depends on the total number of chips in the inverter [25]. Furthermore fault tolerant topologies enable a lifetime extension in case of a failure. To reduce the maximum temperature, another promising field is the cooling system. The design can be made for higher heat dissipation, which is especially possible with heat pipes or Peletier Devices included in the DBC substrate [5]. Another possibility is to control the cooler for avoiding fast temperature changes [26]. A problem with the cooling is the unequal cooling effect of different locations and thus semiconductors of a power module, e. g. water which gets heated up by passing the first semiconductor before cooling the last semiconductor. This problem can be addressed with a suitable structure of the heatsink for the expense of increased production costs of the heatsink or in [27] with a liquid sprayed on the cooled side of the module. This is cost efficient and enables double sided cooling with equal heat distribution.

B) Condition Monitoring

A possibility for the prediction of a failure before its occurrence is real time monitoring of the device. In literature, it is suggested to observe the thermal resistance between junction and case for prediction of solder fatigue [12]. This is implemented in an online technique for the lifetime prediction in [19], which models the consumed lifetime of the module. The same failure can be observed with low order harmonics, which does not require additional measurement equipment [13]. For the prediction of bond wire lift off failures, it is suggested to measure the collector-emitter voltage, which increases before a failure occurs [1].

![Fig 3: Lifetime extension achieved by system design](image)

The detection of the junction temperature is another way to monitor the conditions. Possible ways of monitoring the junction temperature without direct measurement are the measurement of the $dI_{CE}/dt$, the short circuit current or model based approaches [28],[29],[30]. For more detailed stress monitoring, additional sensors for humidity and vibration could be beneficial [14],[31]. The change in the electrical parameters of an IGBT module can also be detected with high-frequency signal excitation. In particular in [32] no additional signal injection is used for the module diagnostic, instead the PWM pattern acted as excitation voltage. In this work, the authors studied the ringing of the resonant circuit composed by the IGBT parasitic capacitors and the circuit inductance. This allowed detecting an increase of the damping factor with aging; for this reason a band-pass filter tuned at the ringing frequency can be used as a diagnostic signal to assess the power module’s conditions.

C) Active thermal controllers
A possibility to increase the module lifetime with software is the design of a controller, which considers the thermal stress. In general there is the possibility to reduce the load current (category 1), to share the power between paralleled devices (category 2), to modify the loss distribution (category 3) or to displace pulses with an intelligent gate driver (category 4). These approaches can be applied independently from each other in different time ranges, which are classified in Fig. 4.

![Diagram of various chances for thermal controller](image)

**Fig 4: Various chances for thermal controller**

Since the losses of semiconductors are temperature dependent, a loss reduction affects not only a reduced junction temperature but also a reduction of the thermal cycles magnitude $\Delta T$. This is doubly beneficial, because a junction temperature reduction results in an increased lifetime [31]. In [33],[34],[35] a thermal controller is investigated, which applies the first category of thermal controllers with a junction temperature $T_j$ limitation and the third category for the adjustment of the PWM (pulse width modulation) carrier frequency to reduce the losses in the device.

This approach is extended in [7] with a fuzzy controller for the detection of power cycling and its reduction. In particular, a filter is employed to detect the magnitude of the thermal cycles, and this information constitutes the feedback signal for two regulators, that dynamically change the switching frequency and the current limit depending on the operating zone (normal, overtemperature, power cycling high/low).

Further controllers from category 3 change the modulation strategy between continuous and 60 degree discontinuous PWM to archive a loss reduction [10][36],[37]. In [38] an active thermal control of an electric drive composed by an IPM motor and a three-phase, 2-level, back-to-back converter is realized by changing the switching frequency and the maximum current limit. Multiple thermal models, for the motor and the power module as well, were implemented and validated with actual measurements. Differently from other solution presented in literature, the torque control is realized in a novel way, where the maximum torque per ampere (MTPA) control is substituted with the maximum efficiency per Nm (MEPNm), where the losses of the devices/motor are taken into account in the computation of the optimum current values for the selected output torque. Basically, the MTPA and MEPNm loci differ because the reference for the direct current is optimized for the minimum iron losses.

For multilevel inverters the modulation scheme can be used to reduce the thermal stress for single components. In [39] the stress distribution for a 3 level neutral point clamped (NPC) inverter under low voltage ride through (LVRT) conditions is studied. During an LVRT, the grid voltage drops to very low levels, and the inverter has to inject 100% of the rated reactive current, according to the E.ON grid code. In this condition, the current is mainly circulating through the anti-parallel diodes of the power module, and this represents a scenario for which the power module is not designed for. As a matter of fact, the excessive diode conduction reduces the module's lifetime. For this reason, a modified space vector modulation (SVM) was used to synthesize the desired grid voltage balancing the current conduction between IGBT and diode. This potential exists also for other multilevel inverters.
In [40] a novel modulation technique is employed to distribute the commutations across multiple H-bridge cells composing a CHB converter. The algorithm counts the commutations of each cell, and tries to synthesize the desired output voltage in a switching cycle giving priority to the H-bridge which totaled fewer commutations. All of this is realized without compromising the quality of the output waveform. In case of parallel power modules, the total power can be shared intelligently between the devices to avoid high temperatures and high temperature differences between the devices (category 2). This has been investigated for DC/DC converters in [41] or in [42] for inverters in a microgrid.

The possibility to affect losses by circulating reactive power in parallel inverters refers to the same category and has been presented in [43]. This approach prevents the cooling of the module by affecting losses and thus reduces the number of thermal cycles in the module. Disadvantages for this approach are the increased losses due to the higher junction temperature. Since the increased junction temperature reduces the lifetime of a module, the application of this approach needs to be considered carefully. The category 4 is applied for minimum pulse elimination and thus switching loss reduction in [44] without reference to thermal control or additional loss generation, which is also possible. In the same category, it is possible to apply active gate control for controlling the semiconductor losses, without affecting the functionality of the device. This has been applied in [45],[46] to balance the current under thermal imbalance and can also be used for active thermal control. In [47], the gate driver has been adapted to the temperature for counteracting the increased losses caused by an increased temperature [47]. In a similar design, it is possible to decrease thermal cycling with this gate driver.

The discussed approaches for active thermal control are listed in Tab. 1 with a description of controlled variables, their basic ideas, prerequisites, complexity and examples for each approach from literature. The complexity depends on the additional variables needed in the controller and the programming effort for realization while the potential is evaluated by estimation of the potential stress reduction during operation.
Tab. 1: Comparison of the existing active thermal control approaches

<table>
<thead>
<tr>
<th>Subject</th>
<th>Category</th>
<th>Controlled variable</th>
<th>Basic idea</th>
<th>Pre-requirements</th>
<th>Complexity</th>
<th>Applied in:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature limitation</td>
<td>1</td>
<td>Inverter current</td>
<td>$T_j$ limitation</td>
<td>Variability of load current</td>
<td>medium</td>
<td>[7][33][35][37]</td>
</tr>
<tr>
<td>Power sharing</td>
<td>2</td>
<td>Inverter current</td>
<td>Stress reduction</td>
<td>Parallel inverters</td>
<td>Low</td>
<td>[41][42]</td>
</tr>
<tr>
<td>Reactive current circulation</td>
<td>2</td>
<td>Inverter current</td>
<td>Heating of the device</td>
<td>Parallel inverters</td>
<td>High</td>
<td>[43]</td>
</tr>
<tr>
<td>DC-link voltage regulation</td>
<td>2</td>
<td>DC-link voltage</td>
<td>Loss minimization/loss distribution among IGBTs and Diodes</td>
<td>Controllable power semiconductors (no Diode bridges)</td>
<td>Medium</td>
<td>[37]</td>
</tr>
<tr>
<td>Discontinuous/Continuous PWM</td>
<td>3</td>
<td>Modulation scheme</td>
<td>Loss optimal modulation scheme</td>
<td>None</td>
<td>Medium</td>
<td>[8][36][37]</td>
</tr>
<tr>
<td>Switching frequency adjustment</td>
<td>3</td>
<td>PWM carrier frequency</td>
<td>Loss adjustment</td>
<td>None</td>
<td>High</td>
<td>[7][8][33][34][35][36][37]</td>
</tr>
<tr>
<td>Advanced modulation schemes</td>
<td>3</td>
<td>Modulation scheme</td>
<td>Stress balancing</td>
<td>Multilevel inverters</td>
<td>High</td>
<td>[39]</td>
</tr>
<tr>
<td>Intelligent gate drivers</td>
<td>4</td>
<td>Pulse position</td>
<td>Delete/create short pulses</td>
<td>None</td>
<td>Medium</td>
<td>[44]</td>
</tr>
<tr>
<td>Active gate control</td>
<td>4</td>
<td>Gate resistor</td>
<td>Loss adjustment</td>
<td>None</td>
<td>High</td>
<td>[45][46][47]</td>
</tr>
</tbody>
</table>

In Fig. 5 four typical mission profiles are presented, which are used to evaluate the potential of active thermal control in different applications. There is the mission profile of an electric vehicle (EV) with an output power of 45 kW, which has highly varying power in a short time, also changing between consumption and generation of energy. The EV is driven by a synchronous machine, which requests reactive power, whereby in the graphic field weakening is neglected. It is assumed that this mission profile is run two times per day. In Fig 5 b) the mission profile of a 3 MW wind turbine with a permanent magnet synchronous machine is shown, obtaining fast load changes in relatively short time. Failures like LVRT are not considered to be standard conditions and thus not included in the mission profile. The mission profile of an inverter in a traction application is shown in Fig. 5 c). The inverter is rated for 500 kW and feeds an induction machine, whereby field weakening is neglected in the profile. The mission profile is characterized by utilization of the full rated power during acceleration and braking, which are causing high junction temperature fluctuations of the single semiconductors. In Fig. 5 d) the mission profile of a solid state transformer (SST) in the distribution grid is shown. The profile is characterized by slow variations of the temperature in one day periods with low consumption in the night and high consumption during the daytime. It could be extended to an active grid, which means that in certain periods, the generation of the power in the grid is higher than the consumption. Depending on the location of the SST, the ratio between active and apparent power $\cos(\varphi)$ can vary between approximately $0.8 < \cos(\varphi) < 1$. 
For the evaluation of the potential of active thermal control only long term thermal cycles are analyzed and short thermal cycles in the range of seconds and faster are not addressed in this work. It is assumed that the junction temperature of the chips is proportional to the absolute value of the output power. This assumption aggregates the effects of increased semiconductor losses because of increased temperature and improved heat transfer capability of silicon at higher temperatures.

A temperature limitation is especially useful for fast varying power like in EV and TA. If the power is slowly varying with also a constant fundamental frequency, like in WT and SST, the potential is lower, but still facilitating the useful protection against destruction caused by overheating. Power sharing as explained for category 2 is only possible if there are parallel modules operating under partial load, which is nowadays not practical in EV, WT and TA. This potential gets more relevant for the SST application, where a redundancy can exist in the DC/AC stage. The potential of reactive current circulation is similar low for EV, WT and TA because before explained reason. However, it is possible to control the EV and the TA not in the MTPA set point, thus to affect additional losses in the inverter caused by a higher current magnitude. The DC-link voltage control, which refers to the same category, obtains high potential for EV and TA because of the fluctuating power and the partial load. For WT it is less suitable with respect to the minimum DC link voltage necessary to feed the power into the grid. Changing the switching frequency either directly or with the modulation scheme offers high potential in all applications, even if it needs to be considered, that sometimes the losses are only shifted from an inverter to the machine or the rectifier. Advanced modulation schemes are promising in abnormal conditions, which can be LVRT for WT or grid faults in SST, while EV and TA are not offering great potential. The intelligent gate driver offers potential in cases of high modulation indices, where it can delete pulses and thus reduce the losses. This is promising in EV and TA, while WT and SST, which are seldom operating under full load, are less suited. Instead, active gate control is promising for all applications, since it enables to control the losses in every point of operation.

To make a comparison possible, it is assumed that for all systems a conservative design is chosen, in which the switching and conduction losses are equal under rated conditions. For a variable switching frequency, the above introduced wind turbine with a permanent magnetized synchronous machine, operating with a switching frequency of 900 Hz is assumed. In the AC/DC stage, is possible to reduce the switching frequency to 400 Hz and in the DC/AC stage to ten times the switching frequency. This enables a total averaged loss reduction to 67% of the switching losses. An electric vehicle usually operates at high switching frequency to improve the performance in field weakening region. The original switching frequency of 20 kHz can be
reduced in certain situations to 4 kHz, which is a reduction to 20% of the switching losses. In an SST, a reduction of the switching frequency from 15 kHz by 70% is feasible. A higher reduction can lead to violations of the current regulations in the distribution grid. In traction applications, a common switching frequency is 2 kHz, which can be reduced up to 1 kHz, leading to a switching loss reduction of 50%.

From the point of the thermal control approach, switching frequency adaption and active gate control seem to be the most general applicable methods with high potential. The drawbacks of these approaches is that there is a lower limit under which the switching frequency cannot be decreased without incurring in unacceptable performance degradation (i.e., output current harmonics), and for active gate control the thermal regulation is actually realized at the expense of efficiency.

IV. Challenges

Most thermal controllers or real time monitoring techniques require the knowledge about the junction temperature. Even if there are various possibilities for the measurement [48], it is not desired to increase the system costs with additional sensors. Therefore observers, thermal models and lookup tables have been applied [7],[8],[33],[34],[35],[36],[37]. These models still have limited accuracy and increase the calculation effort, without having a feedback that can observe changes in the device. A possible change in the thermal resistance can be caused by aging, which is included in the model of [50] by applying junction temperature feedback with low bandwidth measurement. Nevertheless, the design of a precise model with appropriate calculation burden, wide applicability and high bandwidth remains an open field for research. A possibility to add a feedback to the temperature estimation can come from the use of particular observers like the Extended Kalman Filters, where the information of the thermal model is integrated with already available information [50] or with low-cost low-bandwidth temperature sensor.

The online detection of the thermal cycles is another challenge. In [51] various counting strategies for thermal cycles are reviewed. Widely used is the rainflow algorithm, which is modified for online implementation in [52]. Nevertheless different counting procedures model different stress, whereby the temperature gradient is not considered in any of the existing counting modes.

The validation of a lifetime extension achieved with lifetime control is another field of interest. Early research calculated the mean time between failures (MTBF) or the failures in time (FIT) to evaluate the lifetime of a product [53]. But since the validity of theses parameters is limited, current research shifts to Weibull statistics and the mean cumulative function (MCF) [54],[55]. These concepts require testing until failures for an evaluation of the lifetime. Since these tests have to simulate the product in use on the one side and need to be finished in adequate time on the other side, the stress applied in the test has to be chosen carefully.

V. References

interfaces, in high power IGBT modules, using short circuit current, estimation of power modules, transportation, assumption of IGBT:


- Liserre M.: Real time monitoring and active thermal control, Presentation, Aalborg University, Denmark.


