Aalto University
School of Electrical Engineering
Degree Programme in Automation and Electrical Engineering

Aleksi Vulli

# Accelerated ageing and prognostics of silicon carbide power MOSFET

Master's Thesis
Helsinki, September 19, 2018

| | |
|---|---|
| Supervisor: | Professor Jorma Kyyrä |
| Advisor: | Jari Leppäaho, D. Sc. (Tech.) |

Aalto University
School of Electrical Engineering

Aalto University
School of Electrical Engineering
Degree Programme in Automation and Electrical Engineering

ABSTRACT OF MASTER'S THESIS

| | | | |
|---|---|---|---|
| **Author:** | Aleksi Vulli | | |
| **Title:** Accelerated ageing and prognostics of silicon carbide power MOSFET | | | |
| **Date:** | September 19, 2018 | **Pages:** | xi + 98 |
| **Major:** | Control, Robotics and Autonomous Systems | **Code:** | ELEC3025 |
| **Supervisor:** | Professor Jorma Kyyrä | | |
| **Advisor:** | Jari Leppäaho, D. Sc. (Tech.) | | |

The reliability investigations in power semiconductor components have traditionally concentrated on statistical analysis of the failure data in order to set regular maintenance intervals to prevent failures in the field. A more recent discipline, prognostics, in turn attempts to evaluate online the current state-of-health of the device and to predict the remaining useful life by interpreting signals of degradation. The utilization of prognostics is valuable to businesses as it enables addressing the maintenance only to the products close to failure.

In this thesis we studied prognostics from the physics-based perspective in two types of silicon carbide power MOSFETs, in 11 samples in total. The components were aged in a power cycling test system to produce data of the selected failure precursor, drain-source on-state resistance. For the prognostic analysis we developed a kernel-smoothing-based particle filter and applied it to joint state-parameter estimation in a selected sample. The analysis results indicated satisfactory performance considering the estimation of the states and the parameters but revealed significant deficiencies in the prediction performance of the remaining useful life.

Although the work mainly focuses on studying the power MOSFET as a single component it is important to observe it also as a part of a larger entity. Therefore, at the end of the work we propose design principles for a new test system where the power MOSFET operates in a DC-DC converter. The derived precepts are based on the insight of reliability data analysis and prognostics gained during the study.

| | |
|---|---|
| **Keywords:** | accelerated ageing, prognostics, particle filtering, joint state-parameter estimation, kernel smoothing, silicon carbide, power MOSFET, drain-source on-state resistance, condition monitoring, power cycling |
| **Language:** | English |

Aalto-yliopisto
Sähkötekniikan korkeakoulu
Automaation ja sähkötekniikan koulutusohjelma

DIPLOMITYÖN TIIVISTELMÄ

**Tekijä:** Aleksi Vulli

**Työn nimi:**
Piikarbidi-MOSFET:n kiihdytetty ikäännyttäminen ja prognostiikka

| | | | |
|---|---|---|---|
| **Päiväys:** | 19. syyskuuta 2018 | **Sivumäärä:** | xi + 98 |
| **Pääaine:** | Säätötekniikka, robotiikka ja autonomiset järjestelmät | **Koodi:** | ELEC3025 |

**Valvoja:** Professori Jorma Kyyrä

**Ohjaaja:** Jari Leppäaho, TkT


Tehopuolijohdekomponenttien luotettavuustutkimukset ovat perinteisesti keskittyneet vikadatan tilastolliseen analyysiin säännöllisten huoltovälien asettamiseksi, joilla ehkäistään kentällä tapahtuvia vikaantumisia. Prognostiikka on uudempi tiedonala, joka puolestaan pyrkii määrittämään laitteen käytönaikaisen terveydentilan ja ennustamaan jäljellä olevan elinajan tulkitsemalla signaaleja huonemista. Prognostiikan hyödyntäminen on arvokasta liiketoiminnalle, sillä se mahdollistaa huollon kohdistamisen ainostaan niille laitteille, jotka ovat lähellä vikaantumista.

Tässä diplomityössä tutkimme prognostiikkaa fysiikkaan pohjautuvasta näkökulmasta kahdessa erityyppisessä piikarbiditeho-MOSFET:ssa, kokonaisuudessaan 11 näytteessä. Komponentit ikäännytettiin tehosyklaustestissä nielulähdepäälläoloresistanssidatan keräämiseksi, joka valittiin vikaantumisindikaattoriksi. Prognostista analyysia varten kehitimme ydinsilotukseen perustuvan partikkelisuodattimen, jota sovelsimme yhdistettyyn tilaparametriestimointiin valitussa näytteessä. Analyysin tulokset osoittivat tyydyttävää suorituskykyä tilan ja parametrien estimointissa mutta paljastivat merkittäviä puutteita jäljellä olevan eliniän ennustamisessa.

Vaikka työ pääosin keskittyy teho-MOSFET:n tutkimiseen yksittäisenä komponenttina, on tärkeä huomioda se myös osana suurempaa kokonaisuutta. Tämän vuoksi työn lopussa esitetään suunnitteluperiaatteita uutta testausjärjestelmää varten, jossa teho-MOSFET toimii DC-DC -muuntimessa. Johdetut ohjenuorat pohjaavat työn aikana kertyneelle ymmärrykselle luotettavuusdatan analysoinnista ja prognostiikasta.



**Asiasanat:** kiihdytetty ikäännyttäminen, prognostiikka, partikkelisuodatin, yhdistetty tilaparametriestimointi, ydinsilotus, piikarbidi, teho-MOSFET, nielulähdepäälläoloresistanssi, kunnonvalvonta, tehosyklaus

**Kieli:** Englanti

# Acknowledgements

I would like to thank my supervisor Jorma Kyyrä, I highly appreciate your dedication and guidance which has been influential indeed. I would also like to thank my advisor Jari Leppäaho for the given support. You were always free for a talk when I needed it.

Furthermore, I want to thank ABB Drives and my manager Kjell Ingman for providing this opportunity. Kjell, you have been patient and I hope the results of the study will speak for themselves. I also want to express gratitude to my colleagues in the Q&R laboratory, I think you have all supported me with this project in one way or another. Special thanks go to my colleague and dear friend Joni Jormanainen for his excellent work with the software for combining measurement files given by the used test system.

To my family, I want to thank you for the support you have always given me unconditionally. It has helped me to carry on through this work, my studies, and well, everything. To my beloved girlfriend, Saimi, thank you for your support and patience, this has been a long journey but together we made it.

Helsinki, September 19, 2018

Aleksi Vulli

# Abbreviations and Acronyms

| | |
|---|---|
| ABB | Asea Brown Boveri |
| AF | Acceleration Factor |
| ALT | Accelerated Life Testing |
| ASIR | Auxiliary Sampling Importance Resampling |
| ASIR-KS | Auxiliary Sampling Importance Resampling with Kernel Smoothing |
| CCDF | Complementary Cumulative Distribution Function |
| CDF | Cumulative Distribution Function |
| CBM | Condition-Based Maintenance |
| CM | Condition Monitoring |
| CRA | Cumulative Relative Accuracy |
| CSW | Cumulative Sum of Weights |
| CTE | Coefficient of Thermal Expansion |
| DC | Direct Current |
| DUT | Device Under Test |
| EFR | Early Failure Rate |
| EoL | End-of-Life |
| EoP | End-of-Prediction |
| FT | Failure Threshold |
| IFR | Intrinsic Failure Rate |
| IGBT | Insulated Gate Bipolar Transistor |
| MLE | Maximum Likelihood Estimation |
| MOSFET | Metal-Oxide-Semiconductor Field-Effect Transistor |
| PDF | Probability Density Function |
| PF | Particle Filter |
| PH | Prognostic Horizon |
| PHM | Prognostics and Health Management |
| Q&R | Quality & Reliability |
| RA | Relative Accuracy |
| RUL | Remaining Useful Life |

| | |
|---|---|
| SAM | Scanning Acoustic Microscopy |
| Si | Silicon |
| SiC | Silicon Carbide |
| SIR | Sampling Importance Resampling |
| TF | Time-to-Failure |
| TO | Transistor Outline |
| TSEP | Thermo-Sensitive Electrical Parameter |
| VDMOS | Vertically Diffused Metal-Oxide Semiconductor |

# List of Symbols

| | |
|---|---|
| $\alpha$ | Coefficient in $\alpha$-$\lambda$ metric |
| $\alpha^-, \alpha^+$ | Limits in PH and $\alpha$-$\lambda$ metrics |
| $b$ | Empirical parameter in Arrhenius model |
| $\beta$ | Shape parameter in Weibull distribution |
| $c$ | Empirical parameter in Arrhenius model |
| $C_{GS}$, $C_{DS}$, $C_{GD}$ | Power MOSFET parasitic capacitances |
| $C_M$ | Convergence |
| $\gamma$ | Empirical parameter in Coffin-Manson model |
| $d$ | Duty ratio |
| $\delta$ | Empirical parameter in Coffin-Manson model |
| $D_{FW}$ | Free-wheeling diode |
| $\Delta T_j$ | Junction temperature swing |
| $\frac{\Delta T_j}{dt}$ | Slope of the temperature swing |
| $E$ | Energy |
| $\boldsymbol{\epsilon}_k$ | Measurement noise vector at a time instant k |
| $f$ | Probability density function or Frequency |
| $\boldsymbol{f}$ | State transition function (vector form) |
| $F$ | Unreliability function |
| $\boldsymbol{g}$ | Measurement function (vector form) |
| $h$ | Smoothing factor |
| $\eta$ | Scale parameter in Weibull distribution |
| $i_D$ | Drain current |
| $i_{CH}$ | Channel current |
| $i_{GS}$, $i_{DS}$, $i_{GD}$ | Currents through power MOSFET parasitic capacitances |
| $I_L$ | Load current supply |
| $I_{cycle}$ | Cycling current in MicReD® Industrial Power Tester™ 1500A |
| $I_{sense}$ | Measurement current in MicReD® Industrial Power Tester™ 1500A |

| | |
|---|---|
| $j^{(i)}$ | Auxiliary integer variable of a particle $i$ |
| $l$ | Life measure |
| $l_\lambda$ | Set of all the prediction time indices before time $t_\lambda$ |
| $L$ | Likelihood function in MLE |
| $\lambda$ | Time limit in prognostic metrics |
| $\Lambda$ | Logarithmic likelihood in MLE |
| $\boldsymbol{m}_k^{(i)}$ | Kernel location vector of a particle $i$ at a time instant $k$ |
| $\boldsymbol{\mu}_k^{(i)}$ | Reference point vector of a particle $i$ at a time instant $k$ |
| $M$ | Metric |
| $N$ | Cycles-to-failure |
| $n_s$ | Number of particles |
| $\mathcal{N}$ | Normal distribution |
| $P$ | Index of a prediction a time instant |
| $P_{ave}$ | Average power loss |
| $p\left(\boldsymbol{x}_k \mid \boldsymbol{y}_{1:k}\right)$ | Probability density of a random variable $\boldsymbol{x}$ at a time instant $k$ conditional the observations $\boldsymbol{y}_{1:k}$ |
| $p\left(EoL \mid \boldsymbol{y}_{1:P}\right)$ | Predicted EoL distribution at a time instant $P$ conditional the measurements $\boldsymbol{y}_{1:P}$ |
| $p\left(RUL \mid \boldsymbol{y}_{1:P}\right)$ | Predicted RUL distribution at a time instant $P$ conditional the measurements $\boldsymbol{y}_{1:P}$ |
| $\pi$ | Importance density function |
| $R$ | Reliability function or Resistance |
| $R_{DS_{on}}$ | Drain-source on-state resistance |
| $R_G$ | Gate resistor |
| $RUL_k^*$ | True RUL at a time instant $k$ |
| $\overline{RUL}_k$ | Predicted RUL mean at a time instant $k$ |
| $s$ | Suspension |
| $\sigma^2$ | Variance |
| $T$ | Temperature |
| $T_c$ | Case temperature |
| $\hat{t}_{EoL}^{(i)}$ | Estimated end-of-life for particle $i$ |
| $t_{GP}$ | Miller plateau time |
| $T_j$ | Junction temperature |
| $T_{j_{max}}, T_{j_{min}}$ | Maximum and minimum junction temperatures during cycling |
| $\theta$ | Parameter |
| $\boldsymbol{\theta}_k^{(i)}$ | Parameter vector of a particle $i$ at a time instant $k$ |
| $\bar{\boldsymbol{\theta}}_k$ | Mean of the parameter posterior at a time instant $k$ |

| | |
|---|---|
| $u$ | Sampling variable in systematic resampling |
| $U$ | A single random draw in systematic resampling |
| $\mathcal{U}$ | Uniform distribution |
| $\boldsymbol{v}_k$ | Process noise at a time instant $k$ |
| $\boldsymbol{V}_k$ | Variance matrix of the parameter posterior at a time instant $k$ |
| $V_D$ | Drain voltage supply |
| $v_{DS}$ | Drain-source voltage |
| $v_F$ | Body diode forward voltage |
| $V_G$ | Gate voltage supply |
| $v_{GS}$ | Gate-source voltage |
| $V_{GS_{th}}$ | Gate-source threshold voltage |
| $V_{GP}$ | Miller plateau voltage |
| $V_{on}$, $V_{hot}$, $V_{cold}$ | Voltage measurements in MicReD® Industrial Power Tester™ 1500A |
| $\boldsymbol{w}_k$ | Normalized weights at a time instant $k$ |
| $w_k^{(i)}$ | Normalized weight of a particle $i$ at a time instant $k$ |
| $\tilde{w}_k^{(i)}$ | Unnormalized weight of a particle $i$ at a time instant $k$ |
| $W$ | Sum of weights |
| $\boldsymbol{W}_k$ | Variance matrix of the artificial noise |
| $w_f\,(RUL)$ | Weight factor as a function of RUL at the elements of $l_\lambda$ |
| $x_c$, $y_c$ | Coordinates of the center of mass |
| $\boldsymbol{x}_k^{(i)}$ | State vector of a particle $i$ at a time instant $k$ |
| $\xi$ | Magnitude (level) of a general stress |
| $\xi_{use}$ | Use stress level |
| $\xi_{acc}$ | Accelerating stress level |
| $\boldsymbol{\zeta}_k$ | Artificial noise vector at a time instant $k$ |

# Contents

**Acknowledgements** **iv**

**Abbreviations and Acronyms** **v**

**List of Symbols** **vi**

**1 Introduction** **1**

**2 Accelerated ageing** **3**
2.1 Overstressing . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
2.2 Accelerated life testing . . . . . . . . . . . . . . . . . . . . . . . 6
2.2.1 Distribution fundamentals . . . . . . . . . . . . . . . . . 6
2.2.2 Understanding the analysis . . . . . . . . . . . . . . . . . 7
2.2.3 Data format and parameter estimation . . . . . . . . . . 10
2.2.4 Advanced concepts . . . . . . . . . . . . . . . . . . . . . 13

**3 Prognostics** **15**
3.1 Physics-based approaches . . . . . . . . . . . . . . . . . . . . . 17
3.2 Particle filtering . . . . . . . . . . . . . . . . . . . . . . . . . . 19
3.2.1 SIR filter . . . . . . . . . . . . . . . . . . . . . . . . . . 21
3.2.2 Joint state and parameter estimation . . . . . . . . . . . 25
3.3 RUL prediction and prognostics metrics . . . . . . . . . . . . . 31

**4 Power MOSFET** **36**
4.1 Structure and operation . . . . . . . . . . . . . . . . . . . . . . 36
4.1.1 On-state . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
4.1.2 Turn-on and turn-off . . . . . . . . . . . . . . . . . . . . 39
4.1.3 Losses . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
4.1.4 Junction temperature estimation . . . . . . . . . . . . . 44
4.2 Failure mechanisms and condition monitoring . . . . . . . . . . 45
4.2.1 Challenges . . . . . . . . . . . . . . . . . . . . . . . . . . 46

4.2.2 Package-related mechanisms . . . . . . . . . . . . . . . . 47
4.2.3 Gate-oxide degradation . . . . . . . . . . . . . . . . . . . 51

**5 Power cycling test** **54**
5.1 Test system . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
5.1.1 Measurement method . . . . . . . . . . . . . . . . . . . . 56
5.2 Test description . . . . . . . . . . . . . . . . . . . . . . . . . 59
5.2.1 Parametrization . . . . . . . . . . . . . . . . . . . . . . . 59
5.2.2 Group 1 . . . . . . . . . . . . . . . . . . . . . . . . . . . 61
5.2.3 Group 2 . . . . . . . . . . . . . . . . . . . . . . . . . . . 63
5.2.4 Group 3 . . . . . . . . . . . . . . . . . . . . . . . . . . . 64

**6 Prognostic analysis** **67**
6.1 Failure precursor: change in the on-state resistance . . . . . . 67
6.2 Degradation model development . . . . . . . . . . . . . . . . . 69
6.3 Estimation algorithm and metrics . . . . . . . . . . . . . . . . 70
6.4 Results and discussion . . . . . . . . . . . . . . . . . . . . . . 73

**7 Considerations for the new test system** **77**
7.1 Testing principles . . . . . . . . . . . . . . . . . . . . . . . . . 77
7.2 Selected requirements . . . . . . . . . . . . . . . . . . . . . . . 79

**8 Conclusions** **81**

**Bibliography** **83**

**A Power cycling settings and results** **98**

# Chapter 1

# Introduction

Reliability and lifetime management have become central to the development of power semiconductor components. This has been affected by the emergence of new application fields where the power semiconductors are in a more critical position for safety and closer to an everyday consumer. Furthermore, the general trend of electronics' miniaturization and the development of wide-bandgap devices have unveiled new type of failure mechanisms that need to be addressed.

Traditionally the reliability research has focused on statistical analysis of the failure data gathered from accelerated ageing tests during the product development phase. The results of the analysis have then been utilized in setting of regular maintenance intervals to prevent device failures in the field. However, as the failure times depend significantly on the conditions the devices operate in, the setting of these intervals is challenging if the intended usage environments vary. In this case the described strategy might not be able to filter the breakdowns, and on the other hand may result in healthy devices being replaced or repaired unnecessarily.

Prognostics is a relatively novel discipline stemming from this inaccuracy, and targets in providing health information of a device during its actual use. By interpreting signals of degradation, so called failure precursors, prognostics aims to determine the current state-of-health and to predict the future degradation to estimate the remaining useful life (RUL). Depending on the selected approach, algorithms from the field of state estimation, statistics and artificial intelligence are incorporated in the process.

Within power semiconductor components the utilization of prognostics is still quite limited. Majority of the research has focused on investigating the precursor monitoring possibilities, which has been found to be impeded by many difficulties. As collecting of suitable data has already been detected challenging, the investigation of prognostic prediction algorithms has lagged

behind. However, some explorations have been conducted especially with power MOSFETs and insulated-gate bipolar transistors (IGBT).

In this work we studied physics-based prognostics in two types of silicon carbide (SiC) power MOSFETs using drain-source on-state resistance as a failure precursor. For the RUL prediction we developed an algorithm for joint state-parameter estimation by kernel-smoothing-based particle filtering and applied it to one of the measurement trajectories. 11 components in total were aged in an accelerated manner in a power cycling test system to produce analysable degradation data in a reasonable time scale.

There are three objectives in the study. Firstly, we aim to demonstrate the RUL prediction process by the physics-based prognostic approaches using the developed method. Secondly, we want to assess the suitability of the selected failure precursor in the two tested device types. Lastly, the final goal of the thesis is to provide design principles for a new test system where the power MOSFETs would be used as a part of a DC-DC converter. The derived precepts are based on understanding of the traditional reliability data analysis and prognostics, gathered during the study.

The thesis is divided into 7 chapters. The theoretical basis for the work is discussed in the first three. General considerations in reliability and accelerated ageing are given in Chapter 2, together with providing insight into the traditional statistical analysis of the failure data. Chapter 3 explores the framework of prognostics focusing on the physics-based approaches and particle filtering as its tool. In Chapter 4, the power MOSFET is reviewed in general, still emphasizing reliability issues and condition monitoring of a few selected failure precursors. The empirical work of the study is reported in the following three chapters. The power cycling test conducted for the SiC power MOSFETs is documented in Chapter 5. In Chapter 6, the development of the prognostic method and the analysis executed on the power cycling test results are described. The considerations for the new test system are given in Chapter 7. Finally, conclusions are drawn in Chapter 8.

# Chapter 2

# Accelerated ageing

As a device ages, concerns in reliability emerge due to continuing degradation of materials. Eventually, this results in a *failure*, when a critical parameter in the device can no longer fulfil the specified requirement for proper functionality [1, p. 1]. Therefore, it is appropriate to state that the term failure does not only cover pronounced mechanical breakdowns, but also such events as decrease of capacitance under certain safety threshold. These various application dependant definitions are referred to as *failure mechanisms*.

The material degradation, and thus the failures, are tightly connected to the *stress* the device confronts. It is important to notice that in this context the term stress is not restricted only to the standard mechanical stress, but rather comprises a more general definition of any external agent that is able to cause degradation [1, p. 1]. Therefore, various physical quantities such as temperature, humidity, vibration, voltage and pressure are also considered as stresses.

Different stresses involve different failure mechanisms. For example, mechanical stress may cause a terminal of an aluminium electrolytic capacitor to detach, whereas the decrease of capacitance results from electrolyte evaporation associated with thermal and electrical stresses [2]. Moreover, also the magnitude of the stress, i.e. the stress level, and other stress-profile-specific parameters may affect on the occurring failure mechanism. For instance, many typically encountered failures in soldered interconnects are related to temperature variations, but fluctuations with extremely steep slope, the so called thermal shocks, may cause different mehcanims compared to more gradual cycling [3].

A failure occurs as the stress overcomes the degrading strength of the material. Due to tiny differences in the materials' microstructure, which can exist even in nearly identically manufactured products, the degradation does not propagate evenly between units, whereby the devices will not fail at the

same time although subjected to similar stress conditions [1, p. 61]. As a consequence, the failure rate of a device population is typically not constant but takes the form of the *bathtub curve* [1, p. 81], shown in Figure 2.1. During the early use of a product, in the region of early failure rate (EFR), a lot of failures occur due to manufacturing defects caused by materials with extremely low strengths. However, the rate decreases sharply as the weak population of the products quickly fails. In the middle, at the bottom of the curve, lies the region of intrinsic failure rate (IFR), in which the rate is quite low and close to constant. The IFR failures are mostly caused by very small defects slightly reducing the strength of the materials in otherwise decent products. After IFR, the failure rate starts to rise rather quickly in the wear-out region. In this phase even the products with the highest strengths eventually begin to fail.

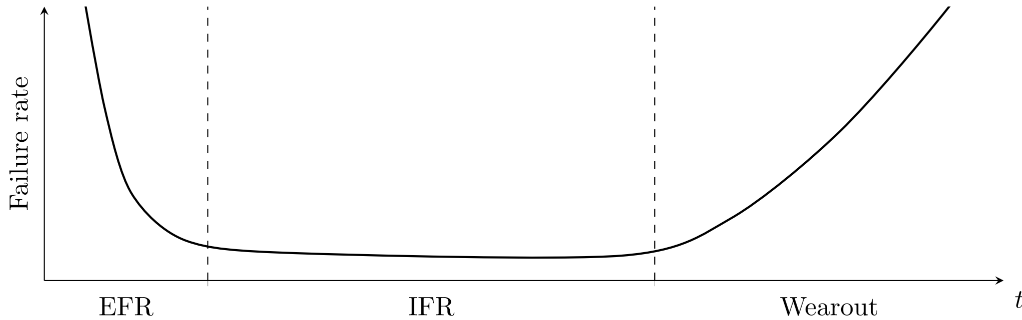



Figure 2.1: Bathtub curve for device reliability.

Tests are required to analyse the occurrence of faults in order to ensure sufficiently reliable operation during the device lifetime. However, since an industrial electronic product may function without failures even for decades if operated within specifications, it may not be possible to base investigations on the *use conditions*. Therefore, to achieve results in a reasonable time scale, *accelerated ageing* is typically the only option.

## 2.1 Overstressing

The accelerated ageing methods can be classified roughly into two: usage-rate and overstress [4]. In the former, the product is operated more frequently than normally, whereby the life of the device is consumed more quickly. However, since the industrial electronic devices considered in this work may function close to continuously, stopped only on occasional maintenance breaks, the latter approach appears more appropriate. In overstress acceleration the

magnitude of a stress attributed to a previously acknowledged critical failure mechanism is elevated to induce failures at a faster rate.

The adjustment of the stress level for accelerated ageing is based on the stress limits of the product, illustrated in Figure 2.2. In the use conditions the device is exposed to a stress that stays within the specification limits. Inside these limits the faultless operation of the product is assumed. If the stress exceeds the specification limits, the product may still work, but is now outside of the zone it is designed to operate in. Moving further, the device ceases to work at the operating limits, but reversion to normal operation is possible provided that the destruct limits are not exceeded. If the applied stress grows higher than the destruct limits the product becomes permanently damaged and is unusable until repaired. Due to the variability in the materials in a device population, the operating and destruct limits are accompanied with distributions to describe the uncertainty around the mean value. The specifications, instead, are strict values set by the manufacturer and therefore do not incorporate any variance.

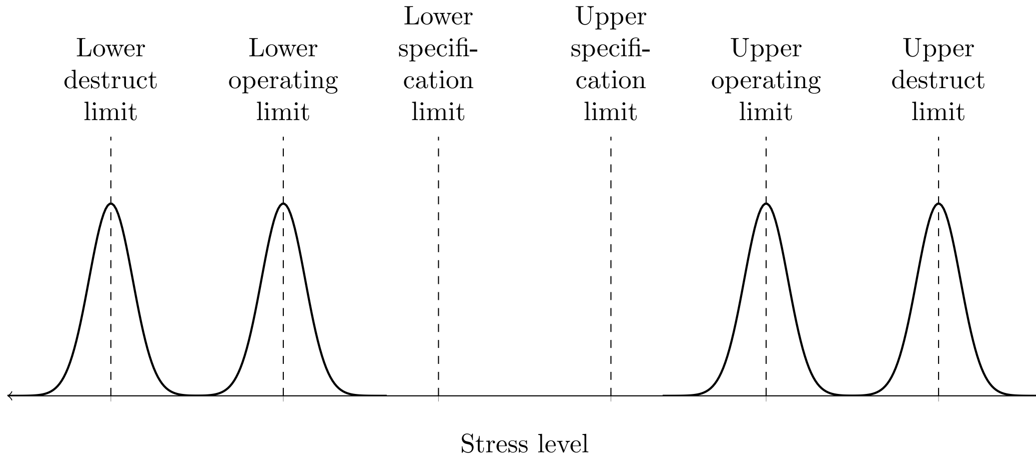



Figure 2.2: An illustration of the device stress limits.

When overstressing is used for accelerating failures for quantitative analysis of a product lifetime, further discussed in 2.2, Vassiliou et al. [4] argue that the stress level should exceed the level the product encounters in its normal use. However, the authors remind that the stresses and the levels should be chosen and adjusted so that they only accelerate the failure mechanisms under consideration, and do not evolve additional mechanisms unnatural in the use conditions. They further note that the higher the stress, the shorter the testing time, but on the other hand as the stress recedes from the use level, the probability of affecting on the physics-of-failure increases. Thus,

the suitable setting for the stress level is typically between the specifications and the operating limits.

## 2.2 Accelerated life testing

As already noted, although exposed under similar environmental and usage conditions, the unit-to-unit variability in the device materials causes the time instants of the failures, the times-to-failure (TF), to distribute. Accelerated life testing (ALT) is a statistical analysis method that exploits the distributing TF data collected from tests in multiple accelerating stress levels, in order to model the life of the devices in the specified use conditions. To ensure sufficient data in a sensible time frame the principles of overstressing are utilized, but as discussed, the stress level has to be adjusted carefully so that the acceleration affects only the failure mechanisms typical in the normal usage. On the other hand, it has to be emphasized that as the ALT analysis is based on the times-to-failure, the process has little to offer if no failures occur.

For business purposes the quantitative reliability analysis is often necessary with many industrial products as it enables estimating the proportion of devices failing before certain time. This knowledge can assist in the specification of service intervals to reduce the number of abrupt maintenance events that tend to be costly. Furthermore, comparison of the ALT results may lead to design improvements, and thus lower the warranty costs and indirectly enhance customer relations via increased product reliability.

### 2.2.1 Distribution fundamentals

To understand the ALT analysis it is inevitable to first consider the basics in statistical distributions. The probability of observing a continuous random variable $T$ within an interval $[a, b]$ is defined as

$$P\left(a \leqslant T \leqslant b\right) = \int_a^b f_T\left(t\right), \tag{2.1}$$

where $f_T\left(t\right)$ is the probability density function (PDF). For simplicity, from hereafter it is marked as $f\left(t\right)$.

From the PDF we can derive the cumulative distribution function (CDF), expressed as

$$F\left(t\right) = \int_{-\infty}^{t} f(x)dx, \tag{2.2}$$

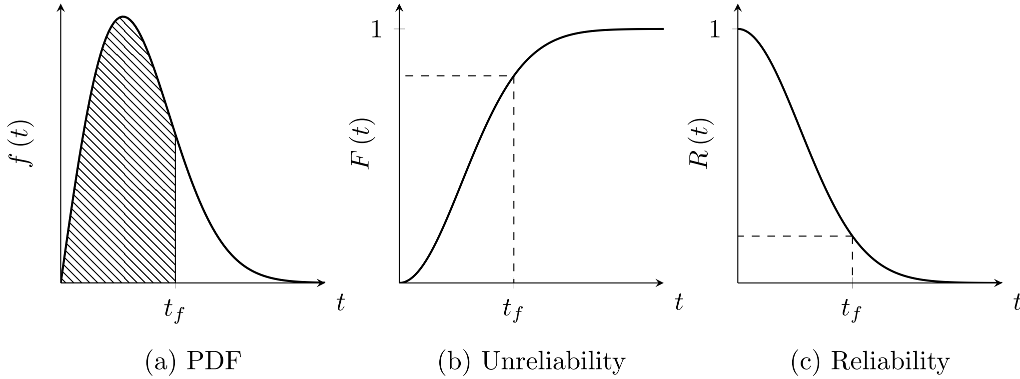
where $x$ is a dummy integration variable. In the context of reliability engineering, the CDF is typically referred to as the unreliability function, as it gives the probability of observing failure in a unit before time $t_f$. Alternatively, the function can be interpreted to describe the proportion of failed units in a population of devices before certain time, here marked $t_f$.

From CDF we can further derive the complementary cumulative distribution function (CCDF) as

$$R(t) = 1 - F(t). \tag{2.3}$$

$R(t)$ is in turn called the reliability function as it describes the probability of *not* observing the failure before $t_f$, or the proportion of healthy units in the population. In Figure 2.3

The functions $f$, $F$ and $R$ are visualized in the form of Weibull distribution in Figures 2.3a, 2.3b and 2.3c, respectively. The probability of observing a failure by time $t_f$ is illustrated with the PDF and the unreliability function. With the reliability function the value at $t_f$ equals the probability of not observing the failure before that time.




(a) PDF (b) Unreliability (c) Reliability

Figure 2.3: The concept of distributions introduced in the form of Weibull.

### 2.2.2 Understanding the analysis

In ALT analysis we search for the dependency between the stress and the distributing times-to-failure of the tested devices. This is defined by conducting the tests in multiple accelerating stress levels in order to model the change of the distribution caused by the varied stress. The constructed model can then be extrapolated to the nominal stress level to reveal the distribution in the use conditions.

The first step in the analysis is to determine the underlying *life distribution* describing the times-to-failure. Typically, it is selected from well-known statistical models based on the following reasons [5]:

1. A physical or statistical argument theoretically supports its connection to a failure mechanism in question
2. The model has already been successfully used with similar failure mechanisms
3. A feasible model provides a decent empirical fit to the data

In relibility engineering the two most frequently used distributions are Weibull and lognormal, of which the former has been especially applied in modelling TF data from semiconductors [1, p. 65]. Therefore, we present the analytical expressions of the PDF and the CDF of the 2-parameter Weibull distribution, defined respectively as

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}\exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right] \tag{2.4}$$

and

$$F(t) = 1 - \exp\left[-\left(\frac{t}{\eta}\right)^{\beta}\right], \tag{2.5}$$

where $\eta$ is the scale parameter and $\beta$ the shape parameter.

In order to model how the life distribution varies by the stress level, a distribution-specific parameter is substituted with a stress-dependent model, *stress-life relationship*. With the 2-parameter Weibull, it is typically addressed to the scale parameter, since the shape of the distribution is considered constant. This assumption of constant $\beta$ implies that the dominant failure mechanism stays the same between different stress levels [4]. However, if the assumption is violated and a good fit is not provided, competing failure mechanisms may have been involved, in which case they typically have to be separated for successfull analysis [6].

The stress-life relationship is often selected from empirically verified models given the accelerating stress. For instance, Arrhenius equation, which originally describes the effect of temperature on the rate of a simple chemical reaction, has been applied especially in modelling temperature acceleration.

However, in statistical analysis it can be generalized for any stress, having the magnitude $\xi$, whereby Arrhenius is defined as [4]

$$l(\xi) = c \exp\left(\frac{b}{\xi}\right), \tag{2.6}$$

where $l$ is a distribution-specific quantifiable life measure and the parameters $b$ and $c$ are reaction specific constants to be defined as a result of the analysis.

Substituting (2.6) for $\eta$ in (2.4) results in

$$f(t,\xi) = \frac{\beta}{c \exp\left(\frac{b}{\xi}\right)} \left[\frac{t}{c \exp\left(\frac{b}{\xi}\right)}\right]^{\beta-1} \exp\left\{-\left[\frac{t}{c \exp\left(\frac{b}{\xi}\right)}\right]^{\beta}\right\}, \tag{2.7}$$

which is the PDF of the Arrhenius-Weibull model. The CDF is derived similarly by substituting the $\eta$ in (2.5) with (2.6) leading to

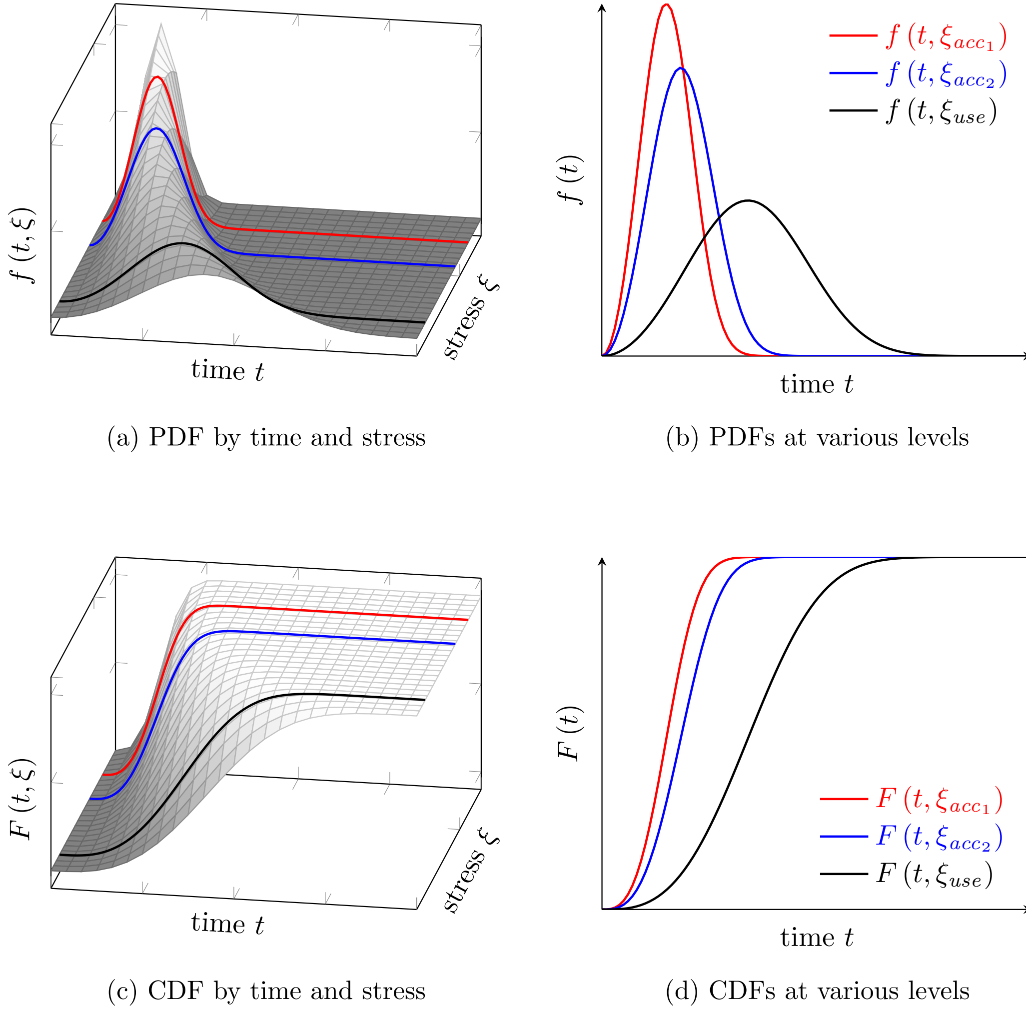
$$F(t,\xi) = 1 - \exp\left\{-\left[\frac{t}{c \exp\left(\frac{b}{\xi}\right)}\right]^{\beta}\right\}. \tag{2.8}$$

Using the methods provided in 2.2.3 the Arrhenius-Weibull model can be fitted to TF data obtained from the tests performed in multiple accelerating stress levels. Testing in two different levels is required in order to fit a single stress model, which is an analogy to the need of two points while assaying a line [4]. It may be possible to construct a more accurate model by testing several levels.

To demonstrate the analysis with two accelerating levels and the resulting extrapolated use level, an example fit of the Arrhenius-Weibull is illustrated. The equations (2.7) and (2.8) with exemplar values of $\beta$, $b$ and $c$ are drawn in respect of time and stress in Figures 2.4a and 2.4c, respectively. Figures 2.4b and 2.4d correspond to the PDF and the CDF with constant $\xi$ in two accelerating levels and in the extrapolated use level.

Another metric typically derived as a part of the ALT analysis is the amount of acceleration produced in the test. This is defined by the *acceleration factor* (AF), describing how much faster the life of a product was consumed in the high stress compared to the use stress. AF is expressed as the ratio of the selected life measure in use and accelerated conditions as

$$AF = \frac{l(\xi_{use})}{l(\xi_{acc})}. \tag{2.9}$$




(a) PDF by time and stress

(b) PDFs at various levels

(c) CDF by time and stress

(d) CDFs at various levels

Figure 2.4: The Arrhenius-Weibull model by time and stress $\xi$, parametrized as $\beta = 3$, $b = 1100$ and $c = 5$.

Thus, for Arrhenius relationship the acceleration factor is defined by combining (2.6) and (2.9), which results in

$$AF = \frac{c \exp\left(\frac{b}{\xi_{use}}\right)}{c \exp\left(\frac{b}{\xi_{acc}}\right)} = \exp\left[b\left(\frac{1}{\xi_{use}} - \frac{1}{\xi_{acc}}\right)\right]. \tag{2.10}$$

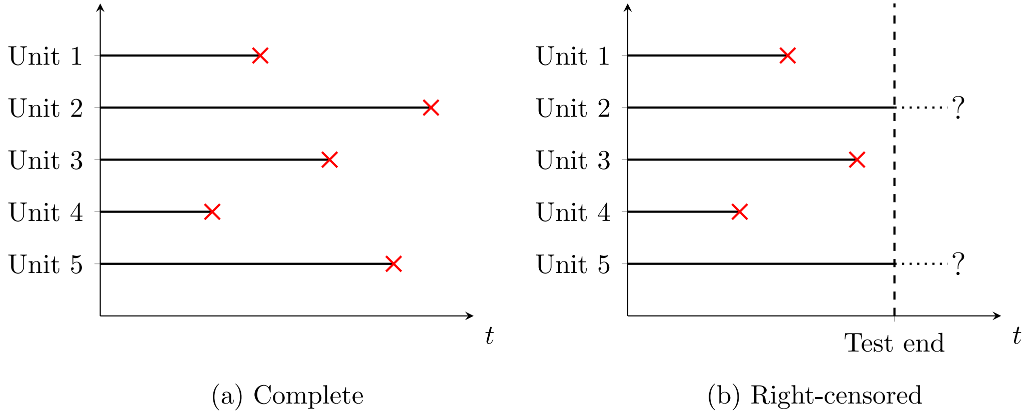
### 2.2.3 Data format and parameter estimation

The ALT analysis culminates in fitting the selected $f(t,\xi)$ to time-to-failure data from various stress levels. Basically, this comprises estimation of the

parameters so that the model describes the collected data as well as possible. To understand the estimation methods to be utilized, the actual data format has to be considered first.

The TF data can be generally divided into *complete* and *censored*. As illustrated in Figure 2.5a, for complete data the exact time-to-failure is known for every unit in the sample. However, due to constraints in the allocated testing time, all of the units may have not failed by the end of the test, or real-time monitoring of each unit may have not been possible. Censored data comprises the observations that can be made in these types of circumstances.

In *right-censoring*, as shown in Figure 2.5b, the exact times-to-failure are known only for some units, whereas others have either survived the test or they have been removed for some reason before a failure. These data are referred to as *suspensions*. Despite the exact TF is not known, it is still assumed that the failures would have occurred eventually. In other words, considering the last observed time instant, the time-to-failure of a unit would lie somewhere to the right on the time-scale. In addition, if the target of the analysis is to model only a single failure mechanism, failures with unrelated mechanisms may also be considered as suspensions [7].




(a) Complete (b) Right-censored

Figure 2.5: The two time-to-failure data formats in the case of 5 units.

To explain the estimation of a parameter vector $\boldsymbol{\theta}$ in a certain model $f\left(t,\xi,\boldsymbol{\theta}\right)$, it is the simplest to deal first only with data in the complete form. Although various methods may be utilized, the most suitable approach in ALT analysis is typically the maximum likelihood estimation (MLE) [4], which searches for such parameter values that maximize the likelihood function. A detailed explanation of the method is omitted here, but the concept of likelihood is elaborated in Chapter 3.

In a time-to-failure data set $t_i$, $i \in 1, 2, ..., n$, each observation is accompanied with the stress level $\xi_i$ the failure occurred in. Here, the likelihood

function becomes [8, p. 174], [9]

$$L(\boldsymbol{\theta} \mid (t_1, \xi_1), ..., (t_n, \xi_n)) = \prod_{i=1}^{n} f(t_i, \xi_i, \boldsymbol{\theta}). \tag{2.11}$$

The distribution parameters, which are now also the maximum likelihood estimators, are therefore obtained by maximizing $L$. Because of the repeated multiplication this is typically quite difficult, whereby the logarithmic version of MLE is often preferred. The logarithmic likelihood function is derived from (2.11) as

$$\Lambda = \ln L = \sum_{i=1}^{n} \ln f(t_i, \xi_i, \boldsymbol{\theta}). \tag{2.12}$$

The estimates which maximize $\Lambda$ may now be found by solving the roots of the partial derivatives as [7]

$$\frac{\partial \Lambda}{\partial \theta_a} = 0 \quad a \in 1, 2, ..., k, \tag{2.13}$$

where $\theta_a$ is a single parameter in the vector $\boldsymbol{\theta}$ of length $k$. The distribution specific analytical expressions do not necessarily exist, in which case numerical methods need to be employed [10, 11].

To include also the right-censored data, (2.11) has to be appended with an additional term to take the aforementioned suspensions $s_j$, $j \in 1, 2, ..., m$ into account. The overall estimation method stays the same whereby the likelihood function is defined as [8, p. 174]

$$L(\boldsymbol{\theta} \mid (t_1, \xi_1), ..., (t_n, \xi_n), (s_1, \xi_1), ..., (s_m, \xi_m)) = \prod_{i=1}^{n} f(t_i, \xi_i, \boldsymbol{\theta}) \times \prod_{j=1}^{m} [1 - F(s_j, \xi_j, \boldsymbol{\theta})], \tag{2.14}$$

from which we can similarly to (2.12) derive the logarithmic form as,

$$\Lambda = \ln L = \sum_{i=1}^{r} \ln f(t_i, \xi_i, \boldsymbol{\theta}) + \sum_{j=1}^{m} \ln [1 - F(s_j, \xi_j, \boldsymbol{\theta})]. \tag{2.15}$$

As previously, the parameters are solved as in (2.13) or with numerical approaches.

### 2.2.4 Advanced concepts

For simplicity, so far we have considered time-to-failure modelling with the assumption of a single stress. In practice, however, the life of an application is often affected by several environmental and usage-dependent stresses. Therefore, to study the product reliability it may be necessary to conduct the tests and the analysis also with multiple simultaneously accelerating stresses.

Also the power semiconductor components considered in this work are susceptible to multiple stresses during their use in various conversion applications. As further discussed in Chapter 4, many failure mechanisms of the power semiconductor packages originate from temperature fluctuations with which the most commonly used stress-life relationship is the Coffin-Manson model [12, p. 86]. It describes the effect of the temperature swing $\Delta T$ on the number of cycles-to-failure based on the inverse power law, expressed as

$$N = \frac{\delta}{(\Delta T)^{\gamma}}, \tag{2.16}$$

where $\delta$ and $\gamma$ are the parameters to be determined as a result of the analysis. In this case, the modelling is based on cycles-to-failure, instead of times-to-failure, which is meaningful as the defining variable is now cycle.

The Coffin-Manson model assumes that the acceleration is attributed solely to the magnitude of the temperature swing. However, empirical evidence has shown that the effect of temperature cycling may depend importantly also on other parameters of the stress profile, whereby Escobar and Meeker [13] introduced a more comprehensive relationship to take into account the cycle maximum temperature $T_{max}$ and the cycling frequency $f$. The former is included as the Arrhenius relationship (2.6) whereas the latter incorporates the inverse power law similarly to (2.16). Thus, the extended Coffin-Manson model becomes

$$N = \frac{\delta}{(\Delta T)^{\gamma_1}} \times \frac{1}{f^{\gamma_2}} \times \exp\left(\frac{b}{T_{max}}\right), \tag{2.17}$$

describing the life of the device by three stresses. As explained in 2.2.2, in the ALT analysis the above model would be substituted for a distribution specific parameter to fit the data utilizing the methods discussed in 2.2.3. Furthermore, the acceleration factor for (2.17) is derived from (2.9) as

$$AF = \left(\frac{\Delta T_{acc}}{\Delta T_{use}}\right)^{\gamma_1} \left(\frac{f_{acc}}{f_{use}}\right)^{\gamma_2} \exp\left[b\left(\frac{1}{T_{max_{use}}} - \frac{1}{T_{max_{acc}}}\right)\right]. \tag{2.18}$$

With multi-stress models the effect of each stress on the life of the device has to be decoupled in order to extrapolate from the accelerating levels to the

use level. Therefore, the tests are conducted with different combinations of the accelerating stress levels. Generally, the more stresses are incorporated, the more testing capacity is required and the more complicated the analysis becomes.

# Chapter 3

# Prognostics

The traditional accelerated life testing discussed in Chapter 2 is primarily used to define suitable maintenance intervals for the devices to be replaced or repaired before a failure takes place. This strategy, referred to as predictive maintenance, has typically been appropriate with productions of large volume when it is feasible to invest resources on the development of the time-to-failure models. However, problems arise when the market of the product covers applications in great variety of environmental and usage conditions. In this case, even if comprehensive time-to-failure models could be achieved by extensive test capacity, the setting of the service intervals becomes challenging due to the widely distributing times-to-failure in the field. This problem can be somehow mitigated through querying the customer of the application characteristics, but the true solution would be some kind of a feedback of the actual health condition of the device.

With some products, thanks to development in sensor technology, the online assessment of the health state has been achieved through condition monitoring (CM) [14]. CM comprises interpreting indicators of impending failures, referred to as *failure precursors*, often seen as a change or a drift in a measurable variable associated with a certain failure mechanism [15]. In power electronic applications this degradation can rarely be measured directly, and estimation from more accessible and sufficiently closely related surrogates is commonly required.

The CM acts a key facilitator in condition-based maintenance (CBM), a strategy to replace or repair only the components that are indicating deterioration and thus a possible failure. As indicated in Figure 3.1, describing the costs associated with the different maintenance strategies, CBM would result in benefits in the most engineering applications [16, p. 1]. In the preventive maintenance the periodical replacements will considerably reduce the number of failures requiring sudden actions, whereby the costs associated with

the repairs are low. However, since all the products are replaced regularly, including also those that would not fail in the near future, the increased reliability is achieved with high prevention costs. In corrective maintenance, in turn, actions are taken only after the failure has occurred, whereby the prevention costs hardly exist. On the other hand, this results in high number of field failures and thus extensive repair costs. Therefore, the cost optimum of the total maintenance, comprising both repairs and prevention, can be found in the region of CBM. Here the number of products requiring repair is moderate, but again, the prevention costs only stem from preventing the failure in those particular products.

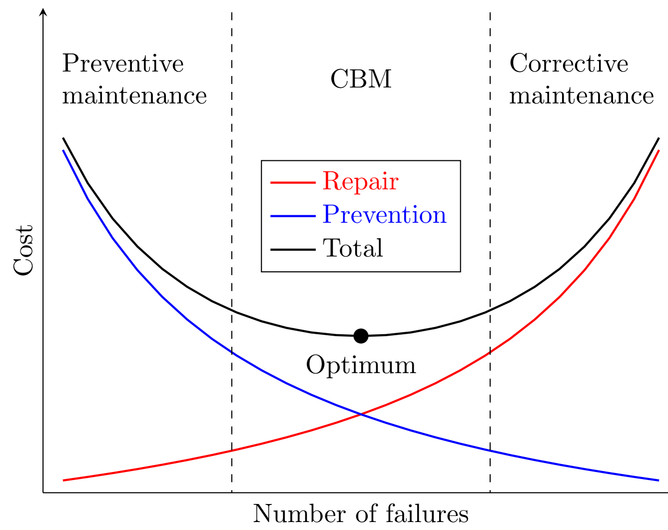



Figure 3.1: Maintenance costs in various strategies as a function of number of occurring failures.

Eventually, the CBM has evolved into a discipline of prognostics and health management (PHM), which can be further divided into two, according to its name. Prognostics focuses on estimating the remaining useful life (RUL) of a device, based on the past and present CM data in combination with the predicted future environmental and operational conditions. Health management in turn utilizes this information in a wider context to perform maintenance, logistics and product life extension intelligently. [17] Originating from the aerospace industry, PHM is a reliatively new field of research, especially applied in mechanical engineering. Within electronics, it has still been utilized only to a limited extent, mainly due to challenges in the condition monitoring of various components [15]. Recently, however, great interest towards prognostics has been developed especially in batteries [18, 19, 20], power semiconductors [21, 22, 23], aluminium electrolytic

capacitors [24, 25, 26] and light-emitting diodes [27, 28, 29].

Because of the novelty of the discipline, the classification of the data analysis methods for RUL prediction still requires disambiguation. The recent grouping into physics-based, data-driven and hybrid approaches [30, 17] seems the most established and should be elaborated. The physics-based methods incorporate a previously identified model whose parameters are estimated based on observed degradation in order to predict the future degradation. The data-driven prognostics in turn involves artificial intelligence or statistical methods to learn the damage progression directly from the measurements, without any actual knowledge of the degradation physics. The hybrid methods pursuit to improve the prediction performance by combining the strengths of the other two approaches. In this study we emphasize the physics-based class, since the majority of the methods investigated under power MOSFET belong into this category. However, as the actual solutions in PHM applications are rarely confined to physics-based or data-driven alone, some hybrid considerations are also included.

## 3.1 Physics-based approaches

Since degradation is associated to natural phenomena, the goal in many studies has been to form models explaining application specific failure precursors. As all models, also these can be divided into physical and empirical. In the framework of prognostics the former category incorporates knowledge on how the degradation is affected by the environmental and usage conditions, whereas the models in the latter class are extracted from numerous laboratory tests by fitting the measured degradation with a mathematical relationship. Generally, supplementing the relationship with knowledge on how the usage conditions affect the degradation is beneficial, but due to challenges in modelling of the degradation physics, empirical models without stress dependencies are still widely used [16, p. 130]. Therefore, the classification of physics-based PHM approaches is sometimes referred to as model-based.

The identified degradation model is utilized in predicting the future behaviour of the degradation in the field use of the device. The parameters of the model can be either derived directly from the laboratory tests, or estimated real-time using the degradation measurement data. The former approach indicates that the model basically depends only on time, and optionally the measured stress conditions, and does not include any feedback of the actual degradation measurement. The latter is based on condition monitoring: the model parameters may be estimated by utilizing degradation

measurements up to the current time instant. As new data is acquired, the estimation process updates the model parameters in pursuit for more accurate and reliable RUL prediction. Due to variability in the device materials and the stress conditions, more precise forecast is achieved if the parameters are updated [30], [16, p. 128].

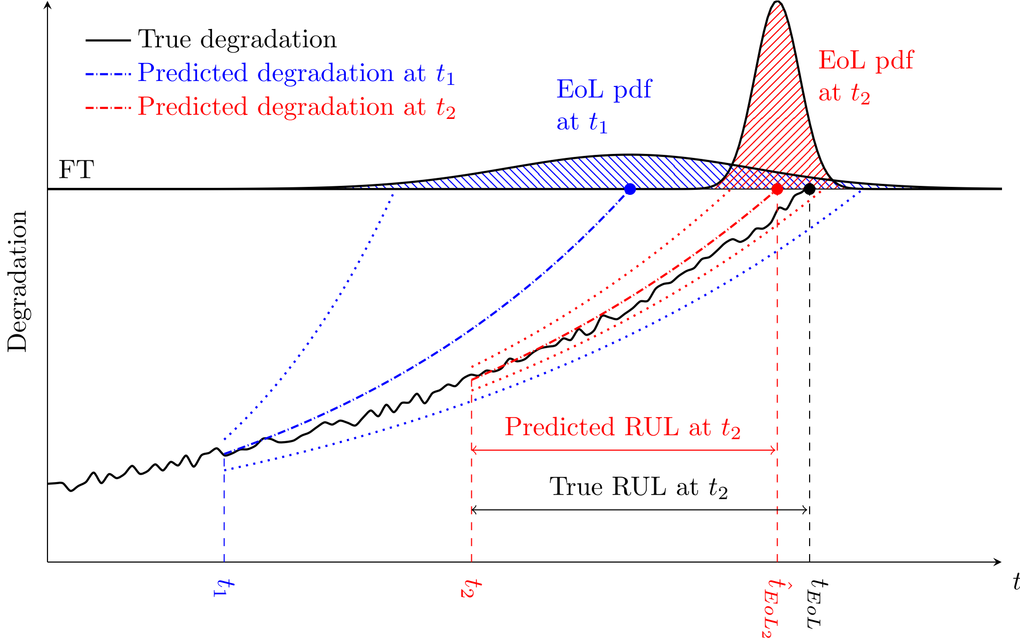



Figure 3.2: RUL prediction process with model parameter updating.

The RUL prediction process in physics-based approaches with model parameter estimation is illustrated in Figure 3.2, where the measured degradation is shown as a black trajectory. At $t_1$ and $t_2$, indicated by the blue and red dashdotted lines, respectively, exponential type of models are extrapolated to the failure threshold (FT), a limit indicating some amount of degradation which is considered intolerable. The model-threshold intersections at $\hat{t}_{EoL_1}$ and $\hat{t}_{EoL_2}$ mark the time instants where the device is predicted to cross the limit, EoL being the abbreviation for end-of-life. The predicted RUL is the difference between $t_i$ and $t_{EoL}$, but due to uncertainty in the data, modelling and existing conditions, the prediction is also associated with some confidence shown as the dotted lines. Therefore, $\hat{t}_{EoL_1}$ and $\hat{t}_{EoL_2}$ are accompanied with distributions $pdf_1$ and $pdf_2$, describing the probability of the prediction. In this particular example, due to longer history of measurements and thus more accurately estimated model parameters, the prediction is more successful at $t_2$ than at $t_1$. This can be noticed from the estimated EoL being closer to the true EoL, and the accompanied PDF being more narrow, which

indicates lower uncertainty. However, both $\hat{t}_{EoL_1}$ and $\hat{t}_{EoL_2}$ still lie inside the corresponding confidences bounds. Considering real applications, it has been proposed that the decision making should be based on the uncertainty measures of RUL rather than the mean value [31].

Considering the physics-based prognostics, several algorithms have been applied in estimation of the degradation model parameters for the RUL prediction. These include for example the Kalman filter [32, 33], the Extended Kalman filter [34, 33], the Unscented Kalman filter [35], the Bayesian method [36] and the particle filter (PF) [34, 18]. In this study however, we focus only on PF, which is is the most commonly exploited framework of algorithms in the field of prognostics, primarily due to its applicability with non-linear processes and non-Gaussian noise [30].

## 3.2 Particle filtering

Particle filtering, or sequential Monte Carlo, is based on Bayesian inference [37] applied for distributions, where observations are utilized in estimating and updating unknown quantities expressed in the form of probability density functions. Fundamentally, the Bayesian inference begins with the prior distribution, which can be interpreted as the probability of a variable before upcoming evidence is taken into account. The evidence, or observation, updates the prior distribution by the likelihood function, the probability of the observation conditional on the prior knowledge. The outcome is the posterior distribution describing the updated probability of the variable conditional on the made observation. The posterior distribution for a random variable $x$ and observation $y$ is defined by the Bayes' theorem as

$$p\left(x \mid y\right) = \frac{p\left(y \mid x\right) p\left(x\right)}{p\left(y\right)} \propto p\left(y \mid x\right) p\left(x\right), \tag{3.1}$$

where $p\left(x\right)$ is the prior distribution of the parameter vector; $p\left(y \mid x\right)$ is the likelihood function, the probability of observing $y$ given the parameters; and $p\left(y\right)$ is considered a normalizing constant as it does not depend on $x$, hence the proportional form is often expressed [16, p. 103]. In all the aforementioned estimation algorithms the Bayesian inference is executed consecutively, referred to as recursive Bayesian update. Thus, the posterior distribution from the previous step is used as the prior distribution in the next step, which in turn is updated whenever a new observation is available.

The significant difference the particle filter has compared to the other aforementioned algorithms, is that the conventionally continuous distributions are approximated with particles, also called samples. In order to de-

scribe the distributions accurately enough with finite number of these particles, PF also incorporates a method called importance sampling, where each particle is weighted based on an arbitrarily selected proposal distribution $\pi\left(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_{1:k}\right)$, also called importance density. Hence, the probability information of the approximated distributions is described by particle locations and their *weights*, as illustrated in Figure 3.3.

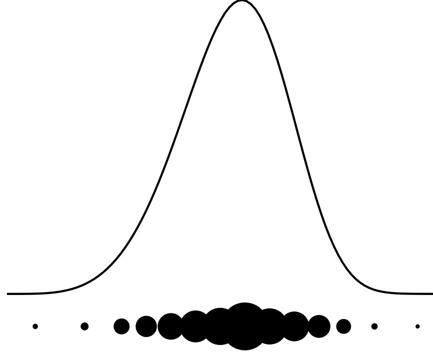


Figure 3.3: An illustration of a distribution approximated by particles, indicated by the circles. The weight of the particle is described by the circle radius.

In sequential importance sampling the weights are corrected based on the recursive Bayesian update. Whenever a new observation is available the *normalized* importance weights $w$ at time instant $k-1$ are updated to the next time instant $k$ by combining the Bayesian inference in (3.1) with the importance density as [38, p. 9]

$$w_k^{(i)} \propto w_{k-1}^{(i)} \frac{p\left(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{y}_k\right)}{\pi\left(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_{1:k}\right)} = w_{k-1}^{(i)} \frac{p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k^{(i)}\right) p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}^{(i)}\right)}{\pi\left(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_{1:k}\right)}, \quad (3.2)$$

where $i$ denotes the index of the particle and $\boldsymbol{y}_{1:k}$ all the past observations.

Unfortunately, due to inevitable increase of variance in the weights, recursive usage of (3.2) results in the degeneracy problem, where after a certain number of iterations all but one particle have negligible normalized weights [39, p. 40]. This phenomenon can be mitigated by employing an additional *resampling* step, where each particle is either eliminated or multiplied based on its weight. In this work we introduce two different PF algorithms exploiting the resampling. The standard sampling importance resampling (SIR) is presented first in 3.2.1 to provide understanding on particle filtering

in general. The auxiliary sampling importance resampling (ASIR), a more complex modification of SIR, is discussed as a part of joint state-parameter estimation in 3.2.2.

### 3.2.1 SIR filter

Sampling importance resampling, an algorithm originally published under the name bootstrap filter [40], incorporates a special case of the importance density which is selected to be the transitional prior. With this modification (3.2) is reduced to

$$\tilde{w}_k^{(i)} \propto p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k^{(i)}\right), \tag{3.3}$$

whereby the weights are expressed completely by the likelihood. The weights are further normalized as

$$w_k^{(i)} = \frac{\tilde{w}_k^{(i)}}{\sum\limits_{j=1}^{n_s} \tilde{w}_k^{(j)}}, \tag{3.4}$$

so that their sum always equals one. Additionally, SIR employs the resampling at every time step.

Due to the easiness in the evaluation of the weights and in the sampling of the importance density, the assumptions required for the usage of SIR can be considered very weak [39, pp. 48–49]. This makes the SIR filter widely applicable in various estimation problems [41], and in the field of prognostics it also seems to be the most employed PF [42].

In order to discuss SIR in detail, we consider state estimation of a system with known and static parameters, a conventional environment for the algorithm to be applied in. Consisting of the state transition function $\boldsymbol{f}$ and the measurement function $\boldsymbol{g}$, a general discrete state-space model is defined as

$$\boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{v}_{k-1}\right) \tag{3.5}$$

$$\boldsymbol{y}_k = \boldsymbol{g}\left(\boldsymbol{x}_k, \boldsymbol{\epsilon}_k\right), \tag{3.6}$$

where $k$ is the index corresponding with the time instant $t_k$, $\boldsymbol{x}_k$ is the state vector, $\boldsymbol{v}_{k-1}$ is the process noise vector, $\boldsymbol{y}_k$ is the output vector and $\boldsymbol{\epsilon}_k$ is the measurement noise vector.

The SIR algorithm can be divided into four steps of initialization, prediction, update and resampling, which are discussed below in detail with the help of Figure 3.4. The illustration considers a section in an estimation process where the SIR filter traverses from the posterior at $k-1$ to the prior at $k+1$.

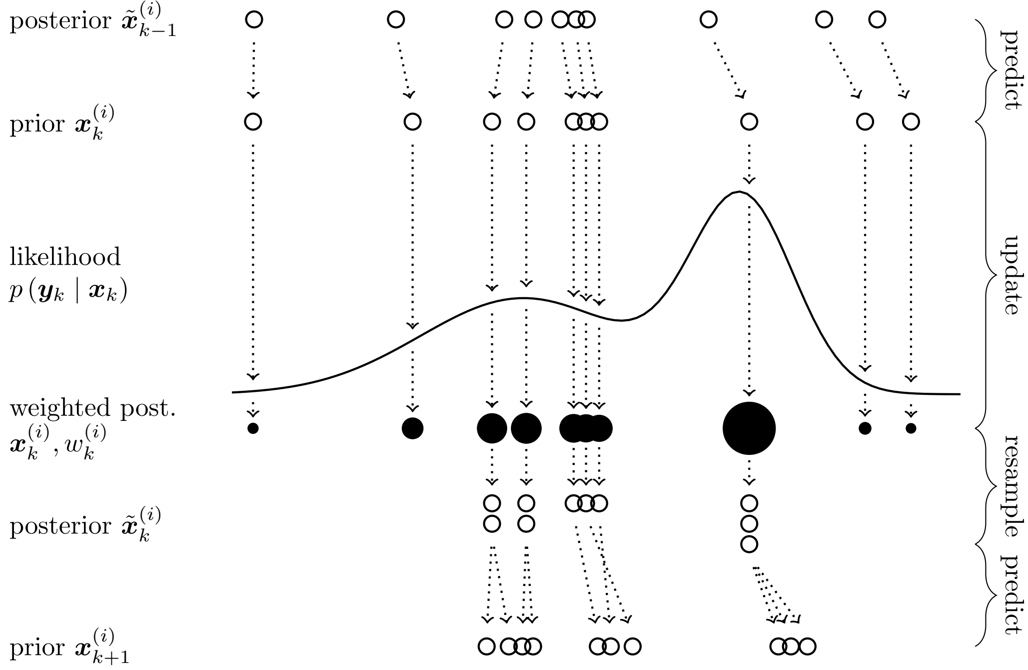



Figure 3.4: Illustration of a section in the sampling importance resampling process, from the posterior at $k-1$ to the prior at $k+1$. The distributions are approximated by the corresponding particles.

**1: Initialization.** The initial distribution is determined and is then approximated by $n_s$ randomly drawn particles $\boldsymbol{x}_0^{(i)}$. The distribution is typically defined based on some initial knowledge of the system state, a common choice is to sample from the uniform distribution [16, p. 154]. In this case the initial particles are drawn as

$$\boldsymbol{x}_0^{(i)} \sim p(\boldsymbol{x}_0) = \mathcal{U}\left(\boldsymbol{r}^-, \boldsymbol{r}^+\right), \tag{3.7}$$

where $\mathcal{U}$ denotes the uniform distribution and its limits $\boldsymbol{r}^-, \boldsymbol{r}^+$ thus characterize the initial state-space.

**2: Prediction.** The particles which approximate the prior distribution as

$$\boldsymbol{x}_k^{(i)} \sim p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}^{(i)}\right) \tag{3.8}$$

are predicted from the initial particles if $k = 1$, or otherwise from the particles approximating the posterior distribution $p(\boldsymbol{x}_{k-1} \mid \boldsymbol{y}_{1:k-1})$. In other words, the posterior particles from $k-1$ are propagated in time to $k$ based on the state transition function (3.5) as

$$\boldsymbol{x}_k^{(i)} = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{v}_{k-1}\right). \tag{3.9}$$

The particles are thus evolved both deterministically and stochastically, by the dynamic model and the process noise, respectively. The latter is also referred to as diffusion, as it separates otherwise overlapping particles.

**3: Update.** The particles of the predicted prior distribution are corrected based on a newly obtained measurement $\boldsymbol{y}_k$. As expressed in (3.3) the likelihood of the observation is calculated for each particle and assigned to the weight. If the measurement noise is assumed normally distributed, i.e. $\boldsymbol{\epsilon} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\sigma}_\epsilon^2\right)$, the likelihood and thus the weights become [16, p. 155]

$$\tilde{w}_k^{(i)} = p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k^{(i)}\right) = \frac{1}{\sqrt{2\pi\boldsymbol{\sigma}_\epsilon^2}} \exp\left\{-\frac{\left[\boldsymbol{y}_k - \boldsymbol{g}\left(\boldsymbol{x}_k^{(i)}\right)\right]^2}{2\boldsymbol{\sigma}_\epsilon^2}\right\}, \tag{3.10}$$

which is the normal probability density having the expected output from (3.6) as the mean. The likelihood is therefore a comparison of the actual observation and the expected measurement, determined by the assumed distribution of the measurement noise. Furthermore, the weights are normalized as in (3.4).

**4: Resampling.** The particles with low weights are likely to be eliminated whereas the particles with high weights are likely to be multiplied. Although various resampling procedures exist [43] only the systematic resampling [44, 45] is considered here due to its computational simplicity and decent empirical performance. The method is based on cumulative sum of weights (CSW), defined as

$$CSW^{(i)} = \sum_{j=1}^{i} w_k^{(j)}, \quad i \in 1, ..., n_s, \tag{3.11}$$

describing the cumulative distribution function of the normalized weights. The weights are then sampled by linking the sampling variables $u^{(a)}$ into $n_s$ evenly spaced sub-intervals as [43]

$$u^{(a)} = \frac{a-1}{n_s} + U \quad a \in 1, ..., n_s, \tag{3.12}$$

where $U$ is a *single* randomly drawn sample from the uniform distribution $\mathcal{U}\left(0, {}^{1}/_{n_s}\right)$. Thus, the sampling variables are also spaced evenly, but the common location in the sub-interval is selected randomly. Then,

by solving for every $a$ the smallest $i$ for which applies $u^{(a)} \leqslant CSW^{(i)}$, the indices of the multiplied particles are defined.

The systematic sampling procedure can also be described with the help of Figure 3.5, where the CSW is generated from 10 weighted particles on the bottom and the systematically set sampling variables are displayed on the left. Starting from the bottom left, the particle associated with the step that first exceeds $u^{(a)}$, is multiplied. The resulting approximated posterior distribution is indicated by the circles on the top. It can be noticed that the probability of resampling a particle with a greater weight is higher compared to a particle with a lower weight.

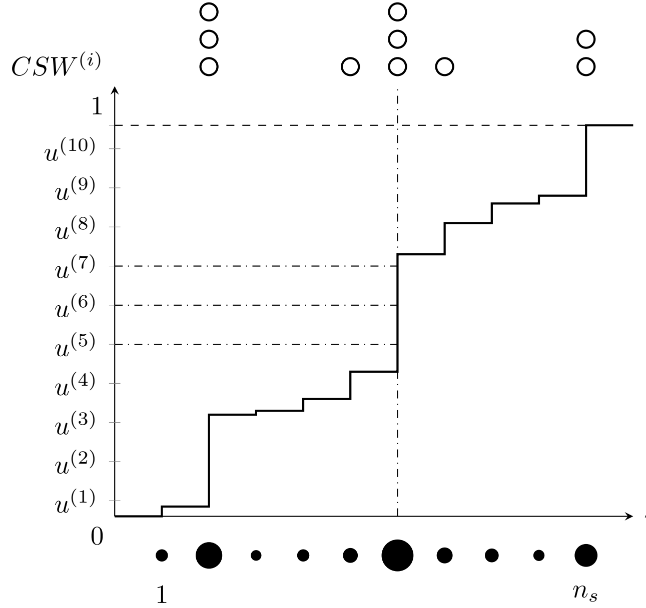



Figure 3.5: Systematic resampling.

In SIR the weights are irrelevant after resampling as they do not depend on the past but only on the likelihood of the current iteration as defined in (3.3). After resampling the next iteration begins from the prediction step.

Based on the above description of the SIR process, Algorithm 1, comprising a single iteration, is presented in pseudo code. It is to be executed whenever a new measurement is obtained. For enhanced modularity, the pseudo code for the systematic resampling is given separately in Algorithm 2.

**Algorithm 1** SIR filter

1: **function** SIR($\boldsymbol{x}_{k-1}, \boldsymbol{y}_k$) ▷ Posterior particles, new meas
2: **for** $i = 1 : n_s$ **do** ▷ For every particle
3: $\boldsymbol{x}_k^{(i)} \leftarrow f\left(\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{v}_{k-1}\right)$ ▷ Prediction: prior
4: $\tilde{w}_k^{(i)} \leftarrow p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k^{(i)}\right)$ ▷ Update: weights from likelihood
5: **end for**
6: $W \leftarrow$ SUM($\tilde{\boldsymbol{w}}_k$) ▷ Sum of weights
7: **for** $i = 1 : n_s$ **do**
8: $w_k^{(i)} \leftarrow \tilde{w}_k^{(i)}/W$ ▷ Normalize weights
9: **end for**
10: $[\boldsymbol{x}_k, -] \leftarrow$ RESAMPLE($\boldsymbol{x}_k, \boldsymbol{w}_k$) ▷ Resample: posterior
11: **return** $\boldsymbol{x}_k$
12: **end function**

**Algorithm 2** Systematic resampling

1: **function** RESAMPLE($\boldsymbol{x}_k, \boldsymbol{w}_k$)
2: $CSW \leftarrow$ CUMSUM($\boldsymbol{w}_k$) ▷ Create CSW
3: $U \sim \mathcal{U}\left[0, 1/n_s\right]$ ▷ Draw the random sample
4: **for** $i = 1 : n_s$ **do**
5: $u^{(i)} \leftarrow (i-1)/n_s + U$ ▷ Create the sampling variables
6: **end for**
7: **for** $i = 1 : n_s$ **do** ▷ For every u
8: **while** $u^{(i)} > CSW^{(j)}$ **do** ▷ Find first level exceeding u
9: $j = j + 1$
10: **end while**
11: $\boldsymbol{x}_k^{(i)} \leftarrow \boldsymbol{x}_k^{(j)}$ ▷ Multiply corresponding particle
12: $j^{(i)} \leftarrow j$ ▷ Get index of the resampled particle
13: **end for**
14: **return** $\boldsymbol{x}_k, \boldsymbol{j}$ ▷ Return resampled particles and indices
15: **end function**

### 3.2.2 Joint state and parameter estimation

So far we have discussed particle filtering in the conventional case of state estimation with fixed and known model parameters. PHM applications, however, are marked by a couple of features opposing this assumption. Firstly, the parameters the degradation model depends on are typically unique to the specific component [46]. Therefore, their value may be assessed only to some extent based on historical data. Secondly, the parameter values may be

further altered from those obtained in laboratory tests due to the environmental conditions appearing in the field [16, p. 153]. Lastly, the parameters may also vary significantly in respect of time, especially if the use conditions are subjected to high dynamics. Therefore, if we consider the degradation as the state, robust RUL prediction often necessitates simultaneous estimation of both the state and the degradation parameters. The challenge we confront falls in the framework of joint state and parameter estimation, for which selected particle filtering methods are discussed below.

The problem is commonly approached by augmenting the parameters $\boldsymbol{\theta}$ into the state vector in the system model [47], in our case (3.5). Thus the combined state-parameter space may be expressed as

$$\boldsymbol{x}_k^* = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{\theta}_k \end{bmatrix}, \tag{3.13}$$

where the parameters can now be considered as state variables, and the problem is reduced to the original state estimation. The Algorithms 1 and 2 presented previously are therefore applicable by replacing $\boldsymbol{x}$ with $\boldsymbol{x}_k^*$. The difference is that each particle now comprises an estimate for both the states and the parameters.

However, in contrast to the states, the model based on which $\boldsymbol{\theta}$ evolve is typically unknown. Nevertheless, the parameters have to be assigned with some type of a transition to enable their estimation. The most simple alternative is the so called persistance model, where the parameters are not changed between the time steps, i.e. $\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1}$ [47, 48]. In this case the transition function (3.5) in combination with (3.13) becomes

$$\boldsymbol{x}_k^* = \begin{bmatrix} \boldsymbol{x}_k \\ \boldsymbol{\theta}_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}_{k-1}, \boldsymbol{v}_{k-1}\right) \\ \boldsymbol{\theta}_{k-1} \end{bmatrix}, \tag{3.14}$$

and we can see that all the possible $\boldsymbol{\theta}$ values for the algorithm to estimate from are set by the initialized particles.

Although the persistance model has its advantages in computational simplicity, it often falls short if the particles have not been initialized properly. More specifically, if the initial distribution is misplaced or the number of particles is not high enough, the estimation performance may suffer as the resampling is to be executed on the same poorly specified set of particles at each time step. The technique appears especially inadequate if the parameters are subjected to significant changes, in which case proper estimation would require a relatively wide initial distribution of $\boldsymbol{\theta}$. However, in the field of prognostics the persistance model has been reported in estimation of the joint state-parameter space in mechanical crack growth [49] as well as power

MOSFET on-state resistance [34]. Here it is essential that the parameters of the degradation model can be defined accurately enough based on some initial or historical measurement data, and that they stay close to static throughout the degradation process.

The assumption of relatively well-known and close to fixed $\boldsymbol{\theta}$ may appear too strict for PHM applications in changing environmental conditions and varied usage. Therefore, we have to consider further methods for estimating unknown and time-variant parameters. A typical approach for applying dynamics into $\boldsymbol{\theta}$ is to incorporate random walk, where an artificial noise term is added into the model to evolve the parameters [47]. In this case, (3.14) is modified as

$$\boldsymbol{x}_k^* = \begin{bmatrix} \boldsymbol{f}\left(\boldsymbol{x}_{k-1}, \boldsymbol{\theta}_{k-1}, \boldsymbol{v}_{k-1}\right) \\ \boldsymbol{\theta}_{k-1} + \boldsymbol{\zeta}_{k-1} \end{bmatrix} \tag{3.15}$$

where the artificial noise vector $\boldsymbol{\zeta}_{k-1}$ may be defined Gaussian as

$$\boldsymbol{\zeta}_{k-1} \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{W}_{k-1}\right), \tag{3.16}$$

for some predefined variance matrix $\boldsymbol{W}_{k-1}$.

Although the random walk has been applied in conventional joint state and parameter estimation techniques, problems arise as the artificial evolution increases the variance of the particles, leading to overdispersed posterior estimates [50, 51]. We may consider a Gaussian parameter posterior, which is approximated by the particles as [51]

$$p\left(\boldsymbol{\theta}_k \mid \boldsymbol{y}_{0:k}\right) \approx \sum_{i=1}^{n_s} w_{k-1}^{(i)} \mathcal{N}\left(\boldsymbol{\theta}_k \mid \boldsymbol{\theta}_{k-1}^{(i)}, \boldsymbol{V}_k\right), \tag{3.17}$$

where $\boldsymbol{V}_k$ denotes the Monte Carlo posterior variance matrix. Now, if we add the artificial noise defined as (3.16), the resulting prior parameter distribution retains the mean $\bar{\boldsymbol{\theta}}_{k+1} = \bar{\boldsymbol{\theta}}_k$, but the variance matrix is increased to

$$\boldsymbol{V}_{k+1} = \boldsymbol{V}_k + \boldsymbol{W}_k, \tag{3.18}$$

which accumulates in respect of $k$.

This inevitable increase of $\boldsymbol{V}_k$ has some unwanted effects. As Daigle and Goebel [46] state, the estimation performance with random walk PF becomes highly sensitive on the selected $\boldsymbol{W}_k$. Generally, large variance yields quick convergence, but the tracking of the trajectory is marked by wide variations. Too small variance in turn, results in very slow convergence, if at all, but in the case of success only small variations occur during the tracking. For prognostics the effect of the random walk is especially harmful

since the accumulated variance propagates into the estimated EoL and RUL distributions unnecessarily increasing uncertainty in the prediction.

To overcome the problem of the increasing variance Liu and West [51] proposed a rather recognized method where kernel smoothing is applied for the parameter posteriors. This is implemented by shrinking $\boldsymbol{\theta}_k$ closer to their mean $\bar{\boldsymbol{\theta}}_k$, to the so called kernel locations expressed as

$$\boldsymbol{m}_k^{(i)} = a\boldsymbol{\theta}_k^{(i)} + (1-a)\,\bar{\boldsymbol{\theta}}_k, \tag{3.19}$$

where $a$ is defined by the smoothing factor $h$ as

$$a = \sqrt{1-h^2}. \tag{3.20}$$

After the shrinkage, a small degree of noise is added with variance $h^2\boldsymbol{V}_k$, whereby the next parameter prior becomes

$$p\left(\boldsymbol{\theta}_{k+1} \mid \boldsymbol{\theta}_k\right) \sim \mathcal{N}\left(\boldsymbol{\theta}_{k+1} \mid \boldsymbol{m}_k^{(i)}, h^2\boldsymbol{V}_k\right), \tag{3.21}$$

which is the evolution assigned for $\boldsymbol{\theta}_k$ in the transition function. Now, the over-dispersion implied by (3.18) is trivially corrected [51, p. 204]. Here we want to remind that as the particles approximate the posterior distribution as in (3.17), their mean and variance matrix are calculated as [52]

$$\bar{\boldsymbol{\theta}}_k = \sum_{i=1}^{n_s} w_k^{(i)}\boldsymbol{\theta}_k^{(i)} \tag{3.22}$$

$$\boldsymbol{V}_k = \sum_{i=1}^{n_s} w_k^{(i)}\left(\boldsymbol{\theta}_k^{(i)} - \bar{\boldsymbol{\theta}}_k\right)\left(\boldsymbol{\theta}_k^{(i)} - \bar{\boldsymbol{\theta}}_k\right)^{\mathsf{T}}, \tag{3.23}$$

where the weights $w_k$ are, as denoted, normalized.

Within particle filtering kernel smoothing is typically incorporated in an auxiliary sampling importance resampling (ASIR) filter [51, 52, 53] originally introduced as a variant of the standard SIR [54]. The fundamental idea in ASIR is to conduct the resampling of the time instant $k-1$ utilizing the knowledge of the measurement at $k$. In this manner the filter pursuits to simulate a situation where the optimal importance density would be available [39, p. 49]. This is achieved with the help of an auxiliary integer variable, denoted as $j^{(i)}$, marking the index of the corresponding particle at $k-1$.

The ASIR filter operates as follows [39, p. 50], [51, p. 201]. The auxiliary integer variables are based on the likelihood of the so called reference points $\boldsymbol{\mu}_k^{(i)}$, which some how characterize the state transition, such as the mean or

a sample. Utilizing the reference points the auxiliary integers are given by resampling the weights

$$w_k^{(i)} \propto p\left(\boldsymbol{y}_k \mid \boldsymbol{\mu}_k^{(i)}\right) w_{k-1}^{(i)}, \tag{3.24}$$

which enable us to draw the actual samples as

$$\boldsymbol{x}_k^{(i)} \sim p\left(\boldsymbol{x}_k \mid \boldsymbol{x}_{k-1}^{j^{(i)}}\right). \tag{3.25}$$

Finally we can derive new weights from the difference in the likelihoods of the actual sample and the reference point as

$$w_k^{(i)} \propto \frac{p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k^{(i)}\right)}{p\left(\boldsymbol{y}_k \mid \boldsymbol{\mu}_k^{j^{(i)}}\right)}. \tag{3.26}$$

As the ASIR is combined with kernel smoothing of the parameter posteriors, the core idea of the filter stays the same with the addition of calculating the kernel locations as in (3.19) and drawing the parameters based on (3.21). Thus, the ASIR-KS filter for a single iteration becomes as in Algorithm 3 [51, 53] presented on the next page. Here the kernel smoothing is performed only once but optionally it can be repeated multiple times [51, 52].

The convergence of the parameter estimates in kernel-density-based particle filters is adjusted by the smoothing factor, for which Chen et al. [55] propose the following settings regarding the required rate of dynamics. If $\boldsymbol{\theta}$ are assumed to be fixed or vary only slowly, a small positive value of $h$ should be set, i.e. $0 < h < 0.2$. On the other hand, if it is expected that the parameters are subjected to significant changes, $h$ should take a value close to 1, i.e. $0.8 < h < 1$. Other methods for defining the smoothing factor include optimization from historical data [55] and a separate online tuning algorithm [56].

Kernel smoothing may be considered as an *adaptive* technique to mitigate the effect of the increasing variance driven by the smoothing factor. In the field of prognostics Hu et al. [57] applied a kernel-smoothing-based PF in the RUL prediction of Lithium-ion batteries. Also other adaptive particle filtering approaches have been proposed to address the variance problem within prognostics. For instance, Daigle and Goebel [46] incorporated SIR with an online $\boldsymbol{W}_k$ tuning algorithm based on relative median absolute deviation of the posterior parameter estimates. The algorithm was applied in RUL prediction of a simulated centrifugal pump with successful results.

In addition to the increase of variance, the state-augemented particle filtering for simultaneous estimation of states and parameters is associated with

**Algorithm 3** ASIR-KS filter

1: **function** ASIR-KS($\boldsymbol{x}_{k-1}, \boldsymbol{\theta}_{k-1}, \boldsymbol{w}_{k-1}, \boldsymbol{y}_k, h$)
2: Calculate $\bar{\boldsymbol{\theta}}_{k-1}$ and $\boldsymbol{V}_{k-1}$ as in (3.22) and (3.23)
3: **for** $i = 1 : n_s$ **do**
4: $\boldsymbol{\mu}_k^{(i)} \leftarrow \boldsymbol{f}\left(\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{\theta}_{k-1}^{(i)}\right)$ ▷ Reference points as the mean
5: $\boldsymbol{m}_{k-1}^{(i)} \leftarrow a\boldsymbol{\theta}_{k-1}^{(i)} + (1-a)\,\bar{\boldsymbol{\theta}}_{k-1}^{(i)}$ ▷ Kernel locations as in (3.19)
6: $\tilde{w}_k^{(i)} \propto p\left(\boldsymbol{y}_k \mid \boldsymbol{\mu}_k^{(i)}, \boldsymbol{m}_{k-1}^{(i)}\right) w_{k-1}^{(i)}$ ▷ Weights
7: **end for**
8: $W \leftarrow$ SUM($\tilde{\boldsymbol{w}}_k$) ▷ Sum of weights
9: **for** $i = 1 : n_s$ **do**
10: $w_k^{(i)} \leftarrow \tilde{w}_k^{(i)}/W$ ▷ Normalize weights
11: **end for**
12: $[-, \boldsymbol{j}] \leftarrow$ RESAMPLE($\boldsymbol{x}_k, \boldsymbol{w}_k$) ▷ Resample for aux integer
13: **for** $i = 1 : n_s$ **do** ▷ Using aux integer do the following
14: $\boldsymbol{\theta}_k^{(i)} \sim \mathcal{N}\left(\boldsymbol{\cdot} \mid \boldsymbol{m}_{k-1}^{j^{(i)}}, h^2\boldsymbol{V}_k\right)$ ▷ Draw new parameters as in (3.21)
15: $\boldsymbol{x}_k^{(i)} \leftarrow \boldsymbol{f}\left(\boldsymbol{x}_{k-1}^{j^{(i)}}, \boldsymbol{\theta}_k^{j^{(i)}}, \boldsymbol{v}_k\right)$ ▷ Predict state by the model
16: $\tilde{w}_k^{(i)} \propto \dfrac{p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_k^{(i)}, \boldsymbol{\theta}_k^{(i)}\right)}{p\left(\boldsymbol{y}_k \mid \boldsymbol{\mu}_k^{j^{(i)}}, \boldsymbol{m}_{k-1}^{j^{(i)}}\right)}$ ▷ Weights as the ratio of likelihoods
17: **end for**
18: $W \leftarrow$ SUM($\tilde{\boldsymbol{w}}_k$) ▷ Sum of weights
19: **for** $i = 1 : n_s$ **do**
20: $w_k^{(i)} \leftarrow \tilde{w}_k^{(i)}/W$ ▷ Normalize weights
21: **end for**
22: **return** $\boldsymbol{x}_k, \boldsymbol{\theta}_k, \boldsymbol{w}_k$
23: **end function**

challenges in computational performance. Inefficiency is observed especially with complex systems as the number of particles required for successful estimation generally increases by the dimension of the joint state-parameter space [46]. As a consequence, different particle filtering approaches for combined state and parameter estimation have also been investigated. These methods include for instance a dual particle filter [41], where two SIR algorithms function in parallel: one is determined to estimate the state and the other the parameters. In a process control application of an ore mill it outperformed the SIR combined with random walk. A similar algorithm was also used in battery state-of-charge estimation with unkown parameters

and the results were successful in comparison with traditional methods [58]. However, as the joint state-parameter space in this study here is considered quite modest, these approaches are not further examined.

## 3.3 RUL prediction and prognostics metrics

Utilizing the methods discussed in 3.2.2, the degradation state and the model parameters are estimated based on the past measurements up to the time instant of prediction $t_k = t_P$. The future behaviour is then predicted by propagating the degradation with the most recently renewed parameter estimates $\boldsymbol{\theta}_P$ to the predominantly set failure threshold. Basically, this equals solving the time $\hat{t}_{EoL}$, the intersection of the predicted degradation trajectory and the FT as

$$d\left(\hat{t}_{EoL}, \boldsymbol{\theta}_P\right) = FT, \tag{3.27}$$

where the function $d$ corresponds with the continuous time-space representation of the discrete state transition function (3.5). However, as the estimated parameters are in the form of a distribution, approximated by particles, also the outcome of the prediction comprises distributions. More specifically, the analysis aims to compute the EoL distribution $p\left(EoL \mid \boldsymbol{y}_{1:P}\right)$, describing the probability density of the times the predicted degradation trajectories and the FT intersect. The RUL distribution $p\left(RUL \mid \boldsymbol{y}_{1:P}\right)$ is then derived by subtracting $t_P$ from $p\left(EoL \mid \boldsymbol{y}_{1:P}\right)$, characterizing the PDF estimate of the time that is left before FT is reached.

The prediction process at $t_P$ is illustrated in Figure 3.6 on the next page and goes as follows. Since there are $n_s$ particles the EoL is computed separately for each $\boldsymbol{\theta}_P^{(i)}$, resulting in $n_s$ possible trajectories from $t_P$ to FT. Thus, there are also $n_s$ estimated EoLs, denoted as $\hat{t}_{EoL}^{(i)}$, which can be derived as the analytical solution of (3.27) given the estimated parameters of each particle. The analytical solution, however, typically exists only with simple models, in which case the EoL distribution may be achieved through simulation. Each particle, having its own estimates of the state and the parameters, is propagated recursively with the transition function (3.5) until the degradation overcomes the failure threshold. Furthermore, as the weight $w_P^{(i)}$ describes the probability of the particle $i$ at $P$, the same weight can be used as the probability of the corresponding $\hat{t}_{EoL}^{(i)}$ [46]. In other words, the most recent weights approximate the predicted EoL probability density at the failure threshold. The predicted RUL distribution is finally calculated by subtracting $t_P$ from each $\hat{t}_{EoL}^{(i)}$.

The RUL prediction performance is of key importance in the development

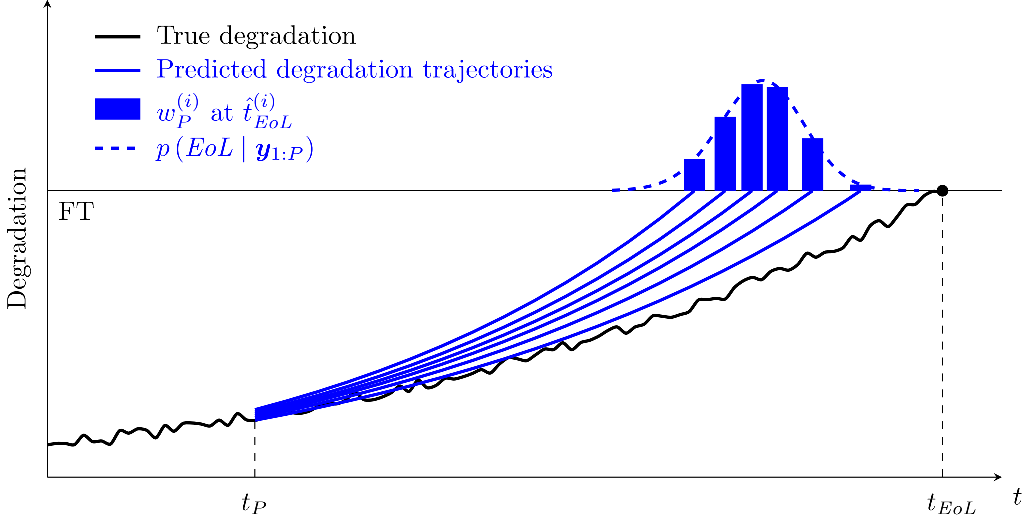



Figure 3.6: An illustration of RUL prediction through simulation at $t_P$. Each particle (here $n_s = 6$) is propagated to the failure threshold, which results in a probability density approximated by the particle weights. The discrete distribution may be smoothed if desired.

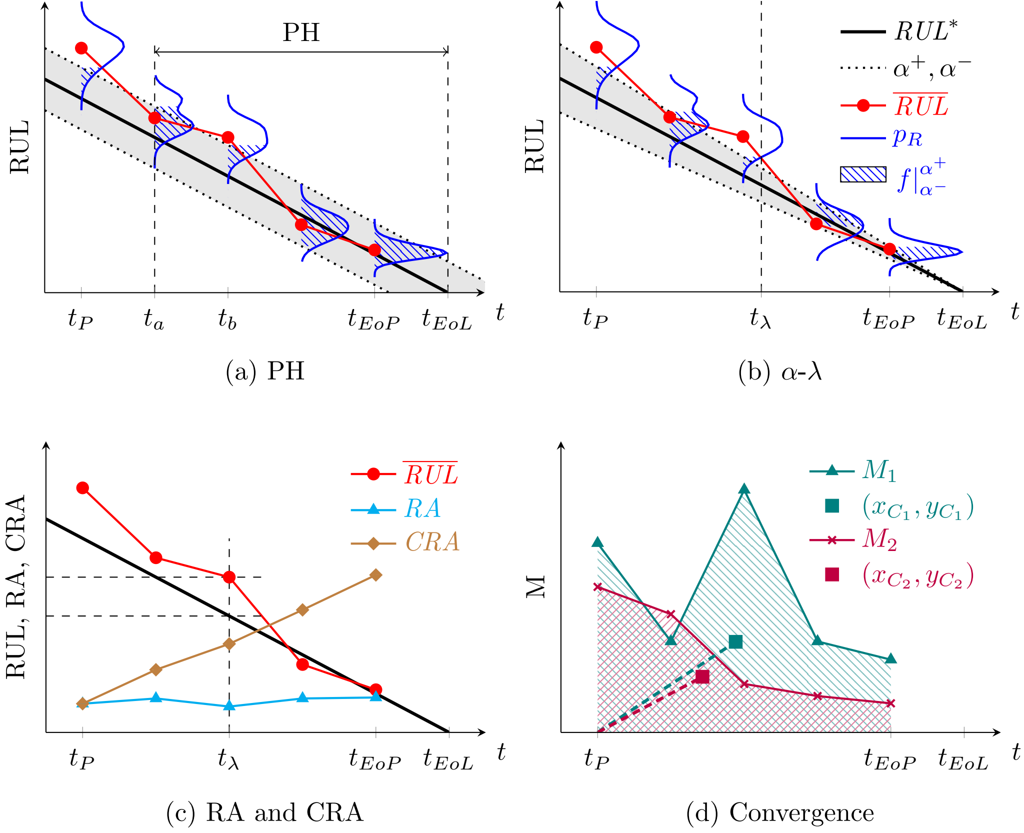
of PHM applications. Therefore, the estimation algorithms are evaluated offline against the true RUL by the true degradation. In these efforts, the degradation measures are commonly obtained from simulations or from accelerated tests performed in laboratory environment. For the performance assessment Saxena et al. [59] propose a process of four tightening metrics including prognostic horizon (PH), $\alpha$-$\lambda$ accuracy, relative accuracy (RA) and convergence. These shall be applied sequentially provided that the requirements set by the previous metrics are fulfilled. Below, we discuss these metrics in detail.

As illustrated in Figure 3.7a, the PH is defined as the time between the EoL and the first time index $a$ when a certain parameter in the predicted RUL distribution, such as the mean $\overline{RUL}$, falls within a fixed accuracy zone specified by the true RUL, which we denote here as $RUL^*$, and the limits $\alpha^+$ and $\alpha^-$. Mathematically the requirement is expressed at each prediction point $j$ as

$$\alpha^- = RUL_j^* - \alpha t_{EoL} \leqslant \overline{RUL}_j \leqslant RUL_j^* + \alpha t_{EoL} = \alpha^+ \tag{3.28}$$

where $\alpha$ is the a coefficient constant describing a certain proportion of $RUL^*$.

As noticed in the graph, in this example the mean later returns beyond the limits at $b$, which may result from sudden unexpected operational transients or other various uncertainties in the actual applications. Although in this case




(a) PH

(b) $\alpha$-$\lambda$

(c) RA and CRA

(d) Convergence

Figure 3.7: The four prognostic performance metrics. The RUL distributions are displayed as smoothed to enhance the visualization. Other alternatives for the visualization of distributions include for instance the box plot or bars describing some characteristic values. Notice also that the $M_1$ in (d) equals $1 - RA$, where $RA$ is from (c), but scaled differently.

an optional approach would comprise a more conservative PH declaration by the last index the $\overline{RUL}$ enters the zone, the authors in [59] encourage to exclude such anomalies by further improving the prediction algorithm and to favour the original definition to avoid confusion. In either method, the longer the horizon is, the earlier it is justified to begin to rely on the forecast. To further emphasize the distributional nature of the prediction process the metric may be improved by regarding the proportion of the probability mass $f\left[p\left(RUL \mid \boldsymbol{y}_{1:j}\right)\right]$ within the accuracy limits. Then the criterion can expressed as

$$f\left[p\left(RUL \mid \boldsymbol{y}_{1:j}\right)\right]\Big|_{\alpha^-}^{\alpha^+} \geqslant \beta \in [0, 1]\,, \tag{3.29}$$

which essentially reverts to (3.28) if $\beta = 0.5$, as the mean marks the half of the probability mass.

The $\alpha$-$\lambda$ increases the accuracy demands in PH by incorporating limits relative to the true RUL at a given time. As seen in Figure 3.7b, since $RUL^*$ approaches to zero, this definition results in a conical accuracy zone determined by the $\alpha$ parameter. The $\lambda$ parameter, in turn, specifies a time instant $t_\lambda$ at which the mean of the predicted RUL distribution must fall within the limits. The $\lambda$ is defined as a proportion of the time scale ranging from the first instant the prediction is applied, to the actual EoL. Thus, the metric employs evaluation of the condition

$$\alpha^- = (1-\alpha)\, RUL^*_\lambda \leqslant \overline{RUL}_\lambda \leqslant (1+\alpha)\, RUL^*_\lambda = \alpha^+, \tag{3.30}$$

where $RUL^*_\lambda$ denotes the true RUL at $\lambda$ and $\overline{RUL}_\lambda$ the corresponding mean of the predicted RUL distribution. Similarly to the PH, also now the accuracy criterion may be set on the probability mass.

In comparison with the $\alpha$-$\lambda$, the RA further quantifies how well the algorithm performs at a given time $t_\lambda$. The metric is defined by the normalized error between the true RUL and the mean of the predicted RUL distribution as

$$RA_\lambda = 1 - \frac{|RUL^*_\lambda - \overline{RUL}_\lambda|}{RUL^*_\lambda}, \tag{3.31}$$

according to which higher RA designates better performance. In order to extend the assessment over a time range, cumulative relative accuracy (CRA) may be used. It is defined as a normalized weighted sum of the evaluated relative accuracies at all prediction points before $t_\lambda$. Thus the CRA is expressed as

$$CRA_\lambda = \frac{1}{|l_\lambda|} \sum_{i=1}^{l_\lambda} w_f\,(RUL)\, RA_\lambda, \tag{3.32}$$

where $l_\lambda$ is the set of all the prediction time indices before $t_\lambda$, $|l_\lambda|$ is the cardinality of $l_\lambda$ and $w_f\,(RUL)$ is a specified weight factor as a function of RUL at the elements of $l_\lambda$. The concept of RA and CRA are illustrated in Figure 3.7c, the latter as a special case where $w_f\,(RUL) = 1$ and $\lambda = EoP$. Furthermore, we set $|l_\lambda| = 1$, i.e. CRA is not normalized, just to clarify the graph.

Finally, the convergence can be regarded as a *meta-metric* as it quantifies how another non-negative precision metric, such as RA, approaches to the optimum. For enhanced conformity it is often desired to consider the error between the instantaneous value and the optimum, whereby the ideal value always equals zero. In this case, the convergence is defined by the euclidean distance between $(t_P, 0)$ and the mass center under the error curve, $(x_c, y_c)$,

and is thus expressed as

$$C_M = \sqrt{\left(x_c - t_P\right)^2 + y_c^2}, \tag{3.33}$$

where

$$x_c = \frac{\frac{1}{2}\sum_{i=P}^{EoP}\left(t_{i+1}^2 - t_i^2\right)M\left(i\right)}{\sum_{i=P}^{EoP}\left(t_{i+1} - t_i\right)M\left(i\right)} \tag{3.34}$$

and

$$y_c = \frac{\frac{1}{2}\sum_{i=P}^{EoP}\left(t_{i+1} - t_i\right)\left[M\left(i\right)\right]^2}{\sum_{i=P}^{EoP}\left(t_{i+1} - t_i\right)M\left(i\right)}, \tag{3.35}$$

where $M\left(i\right)$ denotes the error of the selected metric at a given time index $i$. In Figure 3.7d the convergence is drawn for $M_1 = 1 - RA$, where $RA$ is the one from Figure 3.7c. Additionally, an example metric $M_2$ is illustrated for comparison. The smaller the $C_M$, the faster the convergence, whereby we can notice that $M_2$ performs better.

# Chapter 4

# Power MOSFET

The power MOSFET is widely used as the switch in power electronic converters, especially in the lower power range and in assemblies dedicated for high switching frequency. Furthermore, the development of silicon carbide has enabled even faster operation and usage in high temperatures and voltages. Although this study considers SiC power MOSFETs in the empirical part, due to the novelty of the material we have also included considerations originating from the traditional silicon devices. In this chapter power MOSFET structure, basic operation, failure mechanisms and condition monitoring are discussed.

## 4.1 Structure and operation

The power MOSFET is typically packaged discretely, indicating that it only consists of a single semiconductor chip. The discrete packaging field is today dominated by the transistor outline (TO) family, consisting of many standardized package types of which TO-220 and TO-247 are the most popular representatives. In these outlines, the chip, or the die, is soldered on a copper plate thus forming the die attach. The copper plate is in direct contact with one of the leads while the others are connected to different areas on the chip with aluminium bond wires. All the parts are housed in a transfer mold, leaving bare the ends of the leads and the back side of the copper plate. [60, pp. 348-349] When assembled the leads are soldered to the printed circuit board, and the copper plate is attached to a heat sink. The principle of the package is shown below, on the left side of Figure 4.1, whereas the exterior of a TO-247 package is illustrated on the right.

A traditional power MOSFET die structure used in switched-mode power supplies is Vertical Diffused MOS (VDMOS). "Vertical" stands for mainly

Figure 4.1: Power MOSFET packaging principle. Adapted from [60, p. 349].

upright current flow, whereas "Diffused" refers to the p and n$^{+}$ wells achieved via manufacturing process called diffusing, in this particular case actually double diffusing. Normally, n-channel enhancement type is preferred in power electronics due to electrons being superior over holes in mobility, and since the normally-off feature is often desired [60, p. 285].

A cross section of the chip, a single n-channel VDMOS cell, is presented below in Figure 4.2. The highly doped n$^{+}$-wells, containing many free electrons, are carved in the p-wells containing holes, and the both are connected to the source metallization. The majority of the volume is taken by the lightly doped n$^{-}$-epitaxial layer, which has significantly less free electrons than the n$^{+}$-wells, but is essential in improving the breakdown voltage rating of the component. The lowest layer in the wafer is another highly doped region, the n$^{+}$-substrate, which is joined to the drain metallization.

Figure 4.2: Structure of an N-channel VDMOS cell.

The electrical operation of the VDMOS can also be expressed with the help of Figure 4.2. By increasing the gate-source voltage $v_{GS}$, the holes in

the p-wells are pushed downwards, whereas the electrons accumulate next to the oxide attracted by the positive charges in the gate. As $v_{GS}$ exceeds a certain threshold voltage $V_{GS_{th}}$, the electron concentration in the surface exceeds the concentration of the holes. Thus a channel, an inversion layer, is formed between the n$^+$-wells and the n$^-$-epitaxial layer, creating a route for the electrons through p-wells. If a positive drain-source voltage $v_{DS}$ is applied, the current starts to flow from drain to source. If $v_{GS}$ drops below $v_{GS_{th}}$, the channel is closed and the current flow is ceased. The "verticality" can now be seen clearly: in the on-state the current flows upward from the drain electrode on the bottom, to the source on the top.

### 4.1.1 On-state

In the on-state the power MOSFET is turned on with a positive gate-emitter voltage bias enabling current flow from drain to source. The total on-state resistance $R_{DS_{on}}$ to the current consists of 6 resistances, associated with different parts between the electrodes, considered to be connected in series [60, p. 292]. Five of them are related to the chip: source layer $R_{n+}$, channel $R_{ch}$, accumulation layer $R_a$, n$^-$-epitaxial layer $R_{epi}$ and n$^+$-substrate $R_{sub}$. The sixth term $R_{wcml}$ is related to packaging factors, including resistances associated with bond wires and their connections, metallizations-to-chip interfaces and die attach soldering [32].

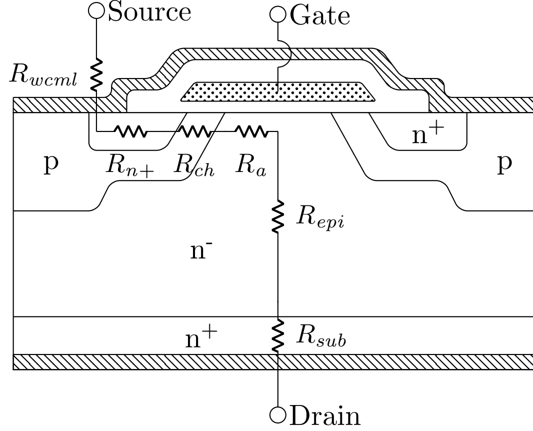



Figure 4.3: N-channel VDMOS with the on-state resistance components.

Thus the total on-state resistance can be calculated as

$$R_{DS_{on}} = R_{wcml} + R_{n+} + R_{ch} + R_a + R_{epi} + R_{sub}, \tag{4.1}$$

Of these, the one related to the n⁻-epitaxial layer has the greatest value in a pristine component. Silicon carbide chips have lower $R_{epi}$ than traditional silicon-based power MOSFETs. [60, p. 309]

### 4.1.2 Turn-on and turn-off

The switching of a power MOSFET is characterised by its parasitic capacitances. Parasitic capacitances exist in electronic components as closely situated conductors in different electric potentials create electric field causing electric charge, thus behaving like a capacitor. In power MOSFET these include the gate-source capacitance ($C_{GS}$), the drain-source capacitance ($C_{DS}$) and the gate-drain capacitance ($C_{GD}$), also referred as input, output and Miller capacitance, respectively [61, p. 409].

As the power converters commonly operate under inductive loads, the power MOSFET switching waveforms are commonly studied based on a clamped inductive circuit. Furthermore, to simplify the analysis the load current is often assumed to stay approximately constant during the process [62], whereby the load inductor is replaced with an ideal current source $I_L$ as shown in Figure 4.4. The clamping, in turn, is provided by the free-wheeling diode $D_{FW}$ connected in parallel with $I_L$. Using this *ideal clamped inductive* arrangement we first describe the turn-on process presenting the waveforms in Figure 4.5. Afterwards, an investigation of the turn-off is conducted, and the waveforms are shown in Figure 4.6. In both analyses the parasitic inductances and diode reverse recovery are neglected.

Caused by the aforementioned parasitic capacitances, the power MOSFET turns on in a four-phase process of turn-on delay, drain current $i_D$ rise, drain-source voltage $v_{DS}$ fall and gate-source voltage $v_{GS}$ rise. Initially, the power MOSFET is in the off-state, in other words, the drain current is zero and $v_{DS}$ equals an externally supplied voltage $V_D$. Moreover, the free-wheeling diode $D_{FW}$ is conducting the external load current $I_L$.

At the beginning of the turn-on delay phase, at $t_0$, the switch S2 is opened and the switch S1 is closed, whereby the voltage source $V_G$ begins to supply the gate with current $i_G$, through the gate resistor $R_G$. Thus, $v_{GS}$ begins to increase as $C_{GS}$ is charged by the main part of the gate current, the displacement current $i_{GS}$. On the contrary, the voltage $v_{DS} - v_{GS}$ over the Miller capacitance is decreasing, whereby $C_{GD}$ is simultaneously discharged with the displacement current $i_{GD}$ [63, p. 36],[62],[64]. It is also worth noting that $i_{GS}$ closes its loop to the gate voltage source right from the source pin, but as the channel is not yet conducting, $i_{GD}$ has to flow back to $V_G$ through the free-wheeling diode [65]. The first phase ends at $t_1$, whereupon $v_{GS}$ overcomes the gate-source threshold voltage $V_{GS_{th}}$, enabling current to flow

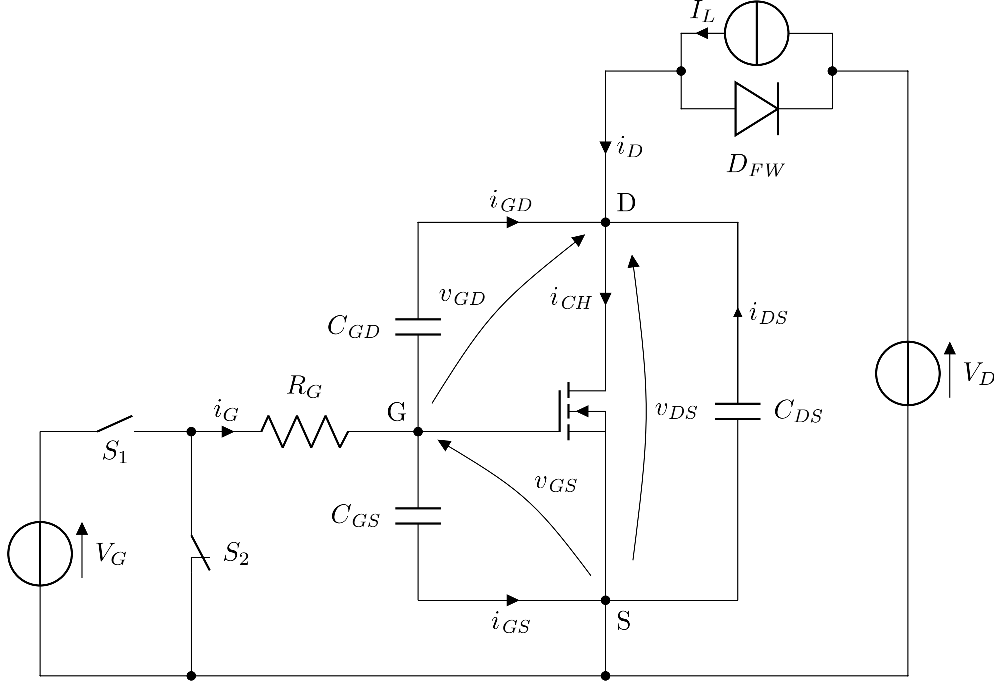



Figure 4.4: Power MOSFET and the parasitic capacitances in an ideal clamped inductive circuit where the load inductor is modelled with an ideal current source $I_L$.

in the channel.

In the $i_D$ rise phase, between $t_1$ and $t_2$, the drain current begins to increase, as the current $i_{GD} + I_L$, which was previously flowing through $D_{FW}$, now starts to change its path towards the opening channel [65]. However, $v_{DS}$ remains constant, as the free-wheeling diode cannot block voltage until all the load current is transferred to the channel. Moreover, $v_{GS}$ continues to increase as $i_{GS}$ is still charging $C_{GS}$.

By the beginning of the $v_{DS}$ fall phase at $t_2$, the drain current has risen up to its maximum $I_L$, and the free-wheeling diode is able to support voltage. Therefore, $v_{DS}$ starts to decrease as the input and Miller capacitances are now discharging through the opened channel, increasing the channel current to $I_L + i_{GD} + i_{DS}$ [63, p. 37],[65]. Although the drain-source voltage at first decreases quickly, the fall is decelerated since $C_{GD}$ and $C_{DS}$ have negative dependence on $v_{DS}$ [63, p. 37],[62]. Eventually, the sign of the voltage over $C_{GD}$ is changed as the decreasing drain-source voltage bypasses the level of $v_{GS}$. Thus, most of the gate current is transferred from the input capacitance to charge the Miller capacitance instead, which has also increased considerably due to negative $v_{DS}$ dependence [63, p. 38]. The drastic reduction in $i_{GS}$

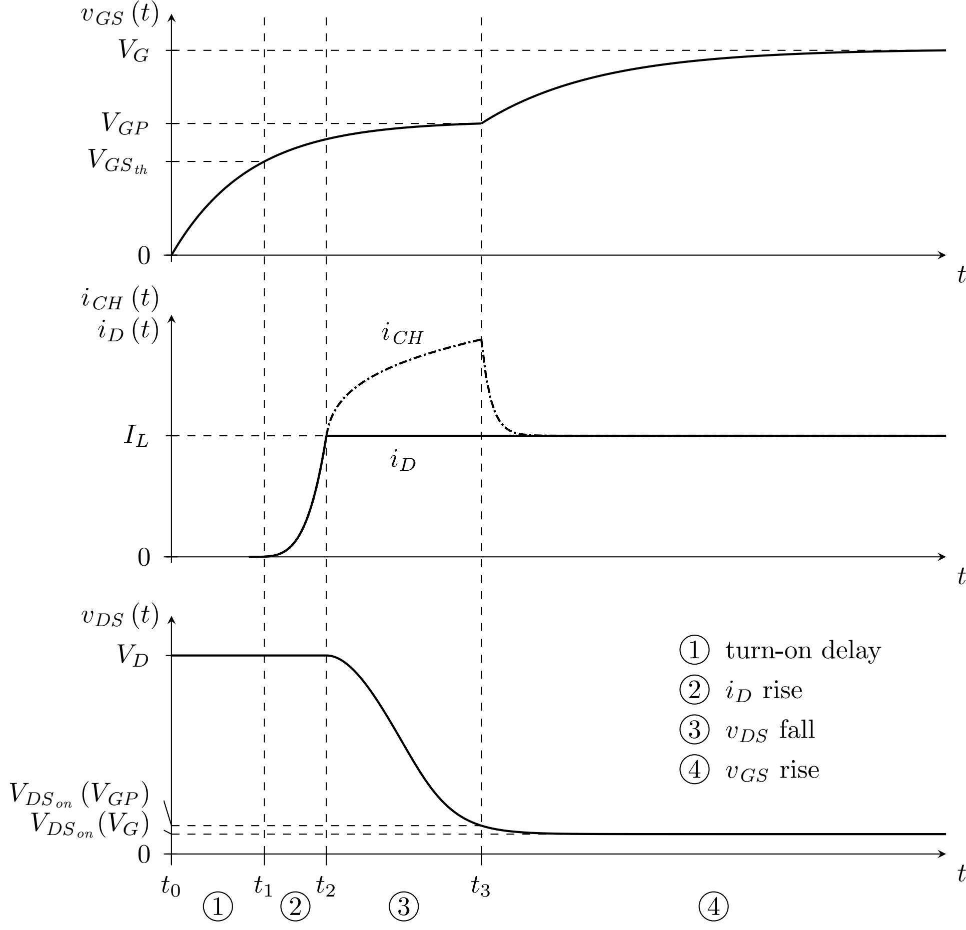



Figure 4.5: The turn-on waveforms of the power MOSFET in the ideal clamped inductive circuit.

results in approximately constant gate-source voltage, a flat region referred as Miller or gate plateau, with an amplitude of $V_{GP}$. The plateau exists until $C_{GD}$ has been charged, during which $v_{DS}$ reduces to the on-state voltage drop at the gate-source voltage of $V_{GP}$, i.e. $v_{DS} = I_L R_{DS_{on}}(V_{GP})$ [61, p. 439],[66]. Although, the drain source voltage actually depends on the channel current, the approximation with the drain current is commonly used, as $i_{CH}$ is not generally measurable.

In the $v_{GS}$ rise phase, starting from $t_3$, the currents $i_{GD}$ and $i_{DS}$ of the charged $C_{GD}$ and discharged $C_{DS}$ become negligible, whereby $i_{CH}$ quickly saturates into its final value $I_L$ [65]. In addition, due to the charged $C_{GD}$, most of the gate current is reverted to charge $C_{GS}$ until $v_{GS}$ reaches the level of $V_G$. The increasing gate-source voltage results in reduction in the on-state

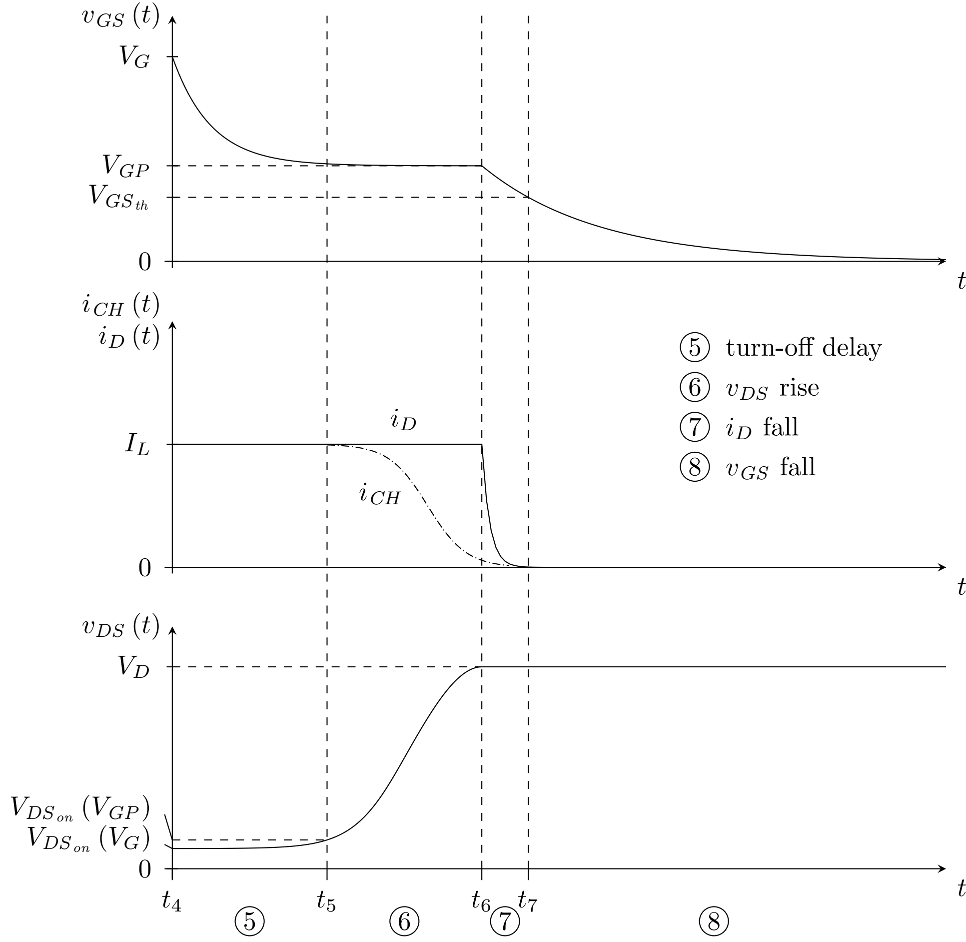



Figure 4.6: The turn-off waveforms of the power MOSFET in the ideal clamped inductive circuit.

resistance, whereby a small decrease in the drain-source voltage down to $I_L R_{DS_{on}}(V_G)$ is yet observed [61, p. 440].

Like the turn-on, also the turn-off process can be divided into four phases: turn-off delay, $v_{DS}$ rise, $i_{DS}$ fall and $v_{GS}$ fall. Initially, the power MOSFET is in the on-state, i.e. $v_{GS} = V_G$, $i_D = I_L$ and $v_{DS} = I_L R_{DS_{on}}(V_G)$.

At the beginning of the turn-off delay phase, at $t_4$, the switch $S_1$ is opened and subsequently the switch $S_2$ closed. This causes the capacitances $C_{GS}$ and $C_{GD}$ to discharge through the gate resistor, whereby the gate-source voltage starts to decrease. However, $I_{DS}$ still remains at its level at $I_L$ as the reverse-biased free-wheeling diode cannot conduct current. In addition, also $v_{DS}$ stays close to constant, although, a small rise is experienced due to increased on-state resistance caused by the decreased $v_{GS}$ [61, p. 440].

In the $v_{DS}$ rise phase, between $t_5$ and $t_6$, the decrease of the drain-source voltage is accelerated due to the negative dependence $C_{GS}$ and $C_{GD}$ possess on $V_{DS}$ [63, p. 39]. Eventually, the voltage over $C_{GD}$ changes sign as $v_{DS}$ bypasses the level of the gate-source voltage, whereby the Miller capacitance starts to charge instead of discharging. Therefore, $i_{GS}$, the current discharging $C_{GS}$, is now compensated by the opposite charging current of $C_{GD}$, and is thus reduced considerably resulting in approximately constant $v_{GS}$ [63, p. 40]. As in the turn-on, this similar flat region in the gate-source voltage in the turn-off process, is called the Miller plateau. In addition, as the drain-source voltage increases, the output capacitance is charged, whereby the channel current is decreased below the drain current [63, p. 39],[67],[68].

By the beginning of the $i_D$ fall phase at $t_6$, the drain-source voltage has risen up to its maximum $V_D$, whereby the charging of the Miller capacitance is finished. The free-wheeling diode becomes forward biased, the load current begins to shift from the channel to the diode and the drain current starts to decrease. The input capacitance is again discharging, whereby the gate-source voltage is also decreased. The drain current reaches the zero level when $v_{GS}$ equals $V_{GS_{th}}$. Finally, in the $v_{GS}$ fall phase, the gate-source voltage is lowered down to zero as the input capacitance is eventually discharged.

### 4.1.3 Losses

The efficiency of a power converter is defined by the losses in the transfer of the electric energy from the input of the converter to its output. The lost energy mostly dissipates into heat, whereby thermal management solutions have to be included and determined depending on the efficiency of the converter. Moreover, in power semiconductor applications, online measuring and calculation of the losses is arranged for junction temperature estimation to be further utilized in diagnostic and prognostic purposes.

In switched-mode converters the total losses are greatly affected by the losses in the switch. For the power MOSFET they can be divided into the operational stages of the on-state, turn-on and turn-off. Technically, a little energy is also lost in the off-state, but since the leakage current in the device is typically in the order of few micro amperes, this stage is generally neglected [60, p. 303]. The total energy losses $E_{tot}$ are thus achieved as the sum of the losses in the three stages, whereby the average power loss $P_{ave}$ in an operational cycle $T$ is expressed as

$$P_{ave} = \frac{E_{tot}}{T} = \frac{1}{T}\left(E_{on} + E_{turn-on} + E_{turn-off}\right), \tag{4.2}$$

where $E_{on}$, $E_{turn-on}$ and $E_{turn-off}$ are the energy losses in the on-state, turn-

on and turn-off, respectively. Furthermore, electric energy is defined as a time integral of the instantaneous electric power given by

$$E = \int_t v_{DS}(t)\, i_D(t)\; dt. \tag{4.3}$$

In reality, in the semiconductor chip the determinative current would actually be the channel current. However, due to practical impossibility of channel current measuring, the approximation to drain current is typically conducted.

In the analysis of the lost power in each of the stages, an ideal clamped inductive circuit in Figure 4.4 from Subsection 4.1.2 is assumed and utilized together with the derived switching waveforms in Figures 4.5 and 4.6. Thus, in the on-state, a static condition is assumed, whereby (4.3) becomes

$$E_{on} = t_{on} V_{DS_{on}} I_{D_{on}} = t_{on} R_{DS_{on}} I_L^2, \tag{4.4}$$

where $t_{on}$ is the duration is the on-state.

As discussed in Subsection 4.1.2, during switching the drain current and drain-source voltage exhibit complicated dependencies on the parasitic capacitances, whereby accurate determination of the power losses is non-trivial. However, these complexities are typically overcome by linear approximation of the rise and fall of $v_{DS}$ and $i_D$, in combination with neglecting the slight changes in the drain-source voltage during the $v_{GS}$ rise phase [61, p. 443]. Therefore, (4.3) can be simplified for turn-on and turn-off as

$$E_{turn-on} = (t_3 - t_1)\, V_D I_L \tag{4.5}$$

$$E_{turn-off} = (t_7 - t_5)\, V_D I_L. \tag{4.6}$$

By substituting (4.4), (4.5) and (4.6) into (4.2), the average power loss for an operational cycle of a power MOSFET is formulated as

$$\begin{aligned} P_{ave} &= \frac{t_{on}}{T} R_{DS_{on}} I_L^2 + \frac{1}{2T}(t_3 - t_1)\, V_D I_L + \frac{1}{2T}(t_7 - t_5)\, V_D I_L \\ &= d R_{DS_{on}} I_L^2 + \frac{f}{2}\left[t_7 + t_3 - (t_5 + t_1)\right] V_D I_L, \end{aligned} \tag{4.7}$$

where $d$ is the duty ratio and $f$ the switching frequency.

### 4.1.4 Junction temperature estimation

Since many of failure mechanisms typically met in power semiconductors originate from temperature variations, as later discussed in 4.2, the estimation

of the temperature inside the component is of great interest. The junction temperature $T_j$, i.e. the temperature in the junction of the semiconductor chip and its solder, is the highest operational temperature in the device and is thus the target of the process. The knowledge of the $T_j$ may be further utilized in diagnostic and prognostic purposes.

The conventional method to derive the junction temperature in power semiconductors exploits the knowledge from the power losses converting into heat. The losses are calculated based on the input power and a thermal impedance model is used to estimate the generated temperature change [69]. The resulted temperature change $\Delta T_j$ is added on top of the simultaneously measured case temperature $T_c$ in order to derive the junction temperature. Unfortunately, with discrete packages, which is the most common package type with power MOSFETs, the sensor is not typically equipped. Furthermore, it has been shown that the device degradation affects the heat path and thus the thermal impedance, whereby updating of the model is required to avoid erroneous estimation [70].

The other possibilities for determining $T_j$ in a power MOSFET can be divided into physical, optical and electrical methods [71]. As the first two require either a physical contact or a visual path to the chip, considering the typically closed package of the power MOSFET we focus only on the electrical methods. These consist of estimating the junction temperature from thermo-sensitive electrical parameters (TSEP), which in case of a power MOSFET include at least the body diode forward voltage, drain-source on-state resistance and gate threshold voltage [70]. Additionally, various parameters in the switching waveforms, such as the delays have been studied in search for new TSEPs [72]. A calibration procedure, where the device is heated in different temperatures for the electrical parameter measurement, is required to decouple the mathematical dependency between the two.

## 4.2 Failure mechanisms and condition monitoring

Failure mechanisms in power semiconductors can be generally classified into two categories: package-related and chip-related [73]. Since the most frequently observed mechanisms in modern power devices originate from the thermomechnical stresses experinced in the semiconductor package [74], they have also acquired the widest attention in the power MOSFET condition monitoring, and are discussed in 4.2.2. As for package-related CM, many studies have also been conducted under IGBTs, which are also applicable

since similar packaging technologies are utilized in both devices. Regarding the chip-related mechanisms, the emerging trend of the silicon carbide devices has brought forth major concerns in the gate oxide reliability which is thus discussed along with possible CM solutions in 4.2.3. However, before further investigating the aforementioned topics in detail, general challenges in condition monitoring of power semiconductors are considered in 4.2.1.

### 4.2.1 Challenges

The adaptation of prognostic approaches into power semiconductor applications is impeded by various difficulties in the condition monitoring. A fundamental problem arises from the failure precursors being often affected by multiple simultaneous degradation mechanisms. This generates uncertainty which has to addressed by the prognostic algorithms and in the worst case may result in a situation where the effects of the mechanisms cancel each other in the followed indicator. On the other hand, if the effect on the measurement is unidirectional it may still be difficult to distinguish what is the actual state of degradation in each of the mechanisms. However, in some cases knowledge on the *total* accumulated degradation may still be sufficient.

Another emblematic issue to the CM in power semiconductors emanates from the temperature dependence of the failure precursors. An observed change may be related to degradation of materials as suspected, or it may be caused by altered junction temperature due to external factors. Thus, conclusions about the health of the device can only be made if the value of $T_j$, at the time instant the degradation measurement is taken, can be measured or estimated.

The most suitable junction temperature estimation method presented for power MOSFET in 4.1.4 was to utilize thermo-sensitive electrical parameters, which are associated with certain difficulties regarding condition monitoring. Also the TSEPs are susceptible to degradation, whereby the temperature measurement may become erroneous as the device ages. The usage of a particular TSEP is thus only valid if it is not significantly affected by the degradation. As an example, Chen et al. [70] utilized gate-source voltage threshold in order to estimate the junction temperature in a power MOSFET assuming that the degradation affects only the die-attach solder and not the actual semiconductor chip. Moreover, many times the TSEPs are the same parameters desired to be condition monitored. In those cases, it has to be taken into account that the same indicator should not be used as both a failure precursor and a temperature sensor [74].

In addition to the junction temperature, many precursors also depend on other variables, such as the drain current in power MOSFET, which is

determined by the normal operation of the converter and thus cannot usually be varied in the means of condition monitoring. As with the junction temperature, in order to derive the health state of the device, the value of $i_D$ has to be known at the time instant the degradation is measured. Generally, a certain calibration step at the commissioning of the converter is required in order to decouple the dependency of the precursor on the affecting variables.

A further issue to be regarded is how to enable condition monitoring in a power semiconductor field application. For a real-online solution the CM should be executed during the normal operation of the converter, complicated by the challenges discussed above. In order to reduce the this complexity some studies have focused on measuring the degradation only on conditions where some of the dependencies can be ignored Dusmez and Akin [75] proposed taking measurements during the starts of the converter when junction temperature may be assumed to be close to ambient. In this case, only a $T_c$ or ambient sensor might be sufficient greatly simplifying the temperature measuring and further facilitating the calibration process. Similarly, Xiong et al. [76] introduced a prognostic application with inverters in electric vehicles where the degradation measurement routine is triggered by key-on and key-off events. Naturally, these approaches require system downtime, but while still being economically and technologically feasible, they might provide enough data for prognostic purposes if a sufficient number of stops during the life of the device can be guaranteed.

Despite the introduced intricacy in the condition monitoring, the most complex task in power semiconductor prognostics may still stem from the connection to the prediction algorithms. Assuming that the CM could be arranged based on some precursor, the physics-based prognostics necessitates knowledge of the critical level of degradation, the failure threshold. Unforunately, we were not able to find any general rule in the literature of how this should be addressed within power devices, but rather the setting of the threshold seems to be highly dependent on the application [74]. It can be expected that a lot ageing tests for a technology in question are required before any generalizations can be made. Even in this case it might be appropriate to assign the threshold with some uncertainty [17], whereby we would consider a distribution of thresholds.

### 4.2.2 Package-related mechanisms

The package-related failure mechanisms in powe semiconductors are mostly attributed to the different coefficients of thermal expansion (CTE) possessed by the various materials of the package and the semiconductor die. Especially the interfaces of joint parts with different CTEs experience stress as

exposed under temperature swings originating from both the application and the environment [77]. The changes in the temperature induced by the active switching of the load current by the semiconductor itself, is typically referred as power cycling, whereas in thermal cycling the swings are caused by alternating ambient temperature.

The greatest temperature variations in power semiconductor packages are generally concentrated on the bond wires and their joints to the chip metallization, originating from the power dissipation in both the die and the wire itself. The generated shear strain in the joint is further increased by a great CTE mismatch between the materials. Repeated flexure of the wire and the shear strain in the joint, most commonly initiate an advancing crack at the tail of the wedge bond, eventually resulting in a complete lift-off of the wire from the bond pad. It is important to note that although the crack is located close to the interface of the wire and the die metallization, it still typically propagates within the wire material, and not at the interface, which in turn would be caused by improper welding or delamination of the metallization. [73] A cross-sectioned bond wire connection with a small initiated crack is shown in Figure 4.7.

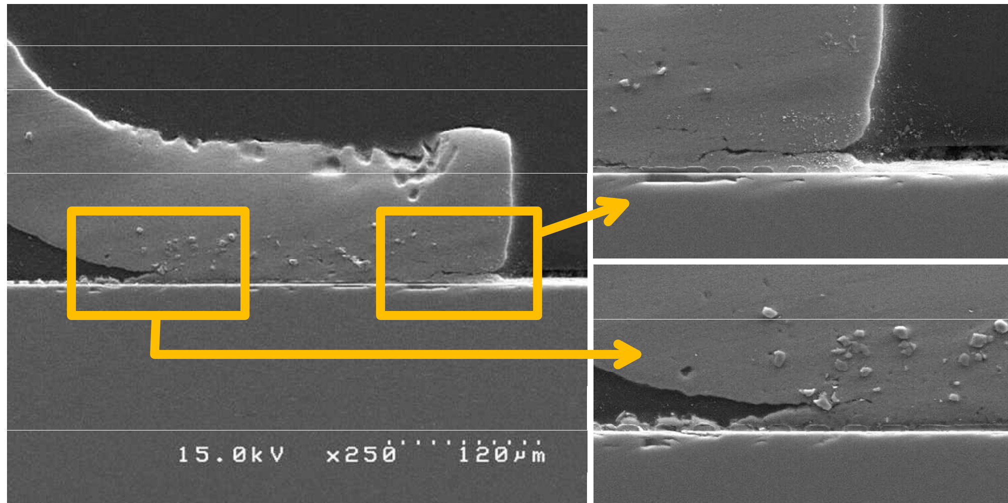



Figure 4.7: A cross-sectioned bond wire connection showing a propagating crack. The image is adapted from [78].

The heel of a bond wire is another possible location for a crack to initiate at, although this failure mechanism is rarely met with advanced power semiconductors. However, heel cracks may occur after long lasting reliability tests, especially if the ultrasonic bonding process has not been properly optimized. As with bond-wire lift-offs, the cause of heel cracking is the repeated flexure of the wire due to temperature variations. [73]

Another critical location in a power MOSFET affected by the temperature variations is the die-attach. Due to the CTE mismatch of the chip, the solder and the copper, shear and tensile strains occur in the material interfaces resulting in deformation of the solder. More specifically, the solder experiences three forms of deformation during a temperature cycle: elastic, plastic and creep [79]. As a consequence, repeated cycling of the temperature induces and augments cracks, which propagate most typically within the solder-copper intermetallic formed in the manufacturing process [73]. The growth of these cracks eventually results in delamination, gradual detachment of the chip and the copper commenced at the corners and at the edges of the die-attach. In addition, the crack growth may create and augment voids, which have been found to propagate towards the center of the die-attach [80]. A scanning acoustic microscopy (SAM) images of two power MOSFET die-attaches showing ageing effects are presented in Figure 4.8. The delamination, indicated by red arrows is much more severe at A, whereas with B only little delamination may be observed. The blue arrow at B points a void.

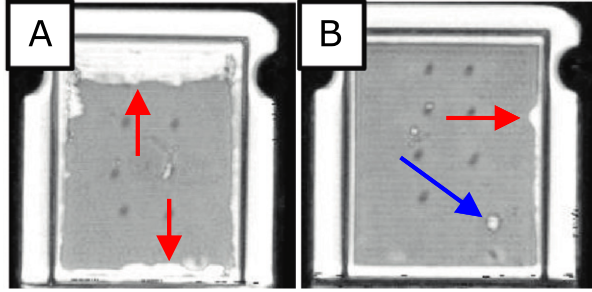



Figure 4.8: SAM images of power MOSFET die-attaches showing delamination (red arrows) and a void (blue arrow). The images are adapted from [81].

A lot of research has been conducted in order to decouple the effect of various properties in temperature variations on the formation of the failure mechanisms discussed above. Generally, the prevalent view in the literature is that the parameter with the most severe effect is the magnitude of the junction temperature swing $\Delta T_j$ [77, 82], but that the mean junction temperature and the slope of the swing $\frac{\Delta T_j}{dt}$ should not be neglected either [83]. More specifically, Schmidt and Scheuermann [82] demonstrated that increasing the duty ratio in a power cycle without altering $\Delta T_j$, accelerates the die-attach ageing but does not particularly influence the bond wire degradation. On the other hand, Dusmez et al. [84] observed die-attach degradation before noticing any significant damage on the bond wires with junction temperature swings of low frequency and relatively high mean temperature. They further

state that with discrete packages it is likely for the solder joint to degrade faster compared to the bond wires when the frequency of the temperature cycling is low.

Considering the condition monitoring of the temperature-related packaging failures in power MOSFETs, one of the most studied precursor is the change in the drain-source on-state resistance [85, 86, 87, 81]. Induced by both the bond wire and the die-attach degradation, the actual mechanism the two affect on $R_{DS_{on}}$ differ from each other. The bond wire cracks and lift-offs reduce the electrical connectivity between the drain and the source, therefore increasing the resistance [88]. The die-attach degradation, in turn, influences by deteriorating the thermal impedance as the heat transfer through the deforming solder is reduced and the temperature dependent $R_{DS_{on}}$ drifts by the junction temperature increase [88, 89, 74]. Additionally, the crack propagation in the solder joint has been shown to increase the electrical resistance as well [90]. Other reported mechanisms affecting $R_{DS_{on}}$ include aluminium reconstruction in the chip metallization [73, 74] and gate oxide degradation [91, 92, 93]. The latter is further discussed separately in 4.2.3, since typically the drain-source on-state resistance has been devoted to studying package-related failures and the gate oxide degradation has been monitored with other indicators.

Because of the numerous presented simultaneously affecting failure mechanisms the power MOSFET condition monitoring by $R_{DS_{on}}$ seems challenging. Consequently, the $R_{DS_{on}}$ trends reported in the literature include some inconsistency. With Si power MOSFETs mainly exponential increase in respect of time or temperature cycles has been observed [85, 34, 94]. In the case of SiC power MOSFETs upward behaviour during cycling tests has also been mostly demonstrated [78, 95], but, on the contrary, Baker et al. [96] reported increase only in the bond wire resistances whereas the total $R_{DS_{on}}$ experienced a slight decrease. The measurement method of the two resistances was based on utilization of the kelvin-source terminal. Additionally, bond-wire lift-offs have been found to induce sudden upturns in the IGBT $V_{CE_{on}}$ measurements [97]. The uncertainty in the observable trend complicates the usage of the drain-source on-state resistance as a prognostic failure precursor.

In addition to the challenges stemming from the various degradation mechanisms, condition monitoring by the drain-source on-state resistance is further complicated by the significant $T_j$ and $i_D$ dependence of $R_{DS_{on}}$. Thus, in order to quantify the degradation $\Delta R_{DS_{on}}$ over time, the relationship $R_{DS_{on}}(i_D, T_j)$ has to be determined in the first place. In the actual applications this so called calibration should be performed during the commissioning, especially if unit-to-unit variability of the power semiconductors

is desired to be neglected. In laboratory environment some integrated solutions already exist, as Stella et al. [98] introduced a method for calibrating a SiC power MOSFET module. The study proposes an application where the temperature of the heat sink is stabilized with external resistors and fans, and the drain-source voltage is measured while injecting a short current pulse into the semiconductor controlling the gate driver. The procedure is then repeated with different pairs of $T_j$ and $i_D$, fundamentally assuming that the current pulse is short enough not to increase the junction temperature significantly, compared to the temperature of the heat sink.

With a calibrated device, possible degradation in the on-state resistance may be quantified by comparing $R_{DS_{on}}$ values derived from the calibrated relationship at specified drain current and with simultaneously estimated junction temperature [75]. However, it has to be noted that also the $R_{DS_{on}}(i_D, T_j)$ relationship experiences change due to aging, whereby recalibration may be required in order to monitor the health condition accurately enough [98].

### 4.2.3 Gate-oxide degradation

The gate-oxide of Si and SiC power MOSFETs consist of silicon dioxide ($SiO_2$) which has been shown to degrade over time in both, although, the actual degradation mechanism differs. With Si, the degradation accumulates positive oxide-trapped charges in the gate oxide, and negative interface-trapped charges at the oxide-silicon interface [99]. On the contrary, with SiC the degradation is caused by direct tunnelling of electrons into near-interface oxide traps, oxygen defects that exist near the SiC-$SiO_2$ interface [100, 101].

Compared to traditional Si devices, the degradation of the gate oxide is much more crucial for SiC power MOSFET. The difference originates mainly from the magnitude of the electric field in the silicon dioxide, whose maximum in SiC power MOSFET is approximately ten times as large as in Si power MOSFET. High electric field induces high threshold voltage, which has to be decreased by reducing the thickness of the gate oxide in order to avoid high $R_{DS_{on}}$. [102] The thinner oxide results in lowered reliability, which has been a major concern preventing the wide usage of SiC power MOSFETs in power electronic applications.

In most studies considering the reliability of the gate oxide, the degradation tests have been conducted under different combinations of constant temperature stress and constant positive gate-source voltage stress. These investigations have shown that under the particular conditions the gate-source threshold voltage drifts due gate oxide degradation in both Si [99] and SiC [100, 101, 103] power MOSFETs. Moreover, similar change has been identified very recently in the Miller plateau voltage and in the Miller

plateau time $t_{GP}$ for both Si [104, 91] and SiC [105] power MOSFETs under positive gate-source voltage stress.

Due to the aformentioned difference in the degradation mechanisms, the time-dependent behaviour of $V_{GS_{th}}$, $V_{GP}$ and $t_{GP}$ observed in the literature show some difference between Si and SiC power MOSFETs. For silicon devices, it has been proposed that over time all the three precursors exhibit a dip-and-rebound variation [104, 91], illustrated in Figure 4.9a. In this pattern a short decrease phase at the beginning, caused by the oxide trapped charges, is soon compensated by the accumulating interface trapped charges. This results in a continuing increase after the turn-around point, eventually exceeding the level of the precursor in a pristine component. On the contrary, with silicon carbide components the dip effect has not been observed, but an increasing trend depending positively on the magnitude and duration of the temperature and gate-bias stress [105]. Moreover, Santini et al. [103] reported that with $V_{GS_{th}}$ the trend is initially steep, but quickly decelerates into more gradual rising, as presented in Figure 4.9c. The effect of the gate-oxide degradation as seen in the gate-source voltage curve during turn-on, is shown for both Si and SiC in Figures 4.9b and 4.9d, respectively.

Constant temperature and gate-source voltage may not resemble very realistically the conditions the power MOSFET encounters as part of a switched-mode DC-DC converter, whereby some studies have considered the gate oxide degradation under more variable stresses. Dusmez et al. [84] exhibited increase in the gate-source voltage threshold of normally switching Si power MOSFETs operating under thermal cycling. Moreover, Fayyaz and Castellazzi [106] reported rising in $V_{GS_{th}}$ of SiC power MOSFETs with pulsed gate but without any load in constant high temperature. On the contrary, Ouaida et al. [107] were not able to observe significant change in $V_{GS_{th}}$ of SiC power MOSFETs under switching conditions in constant high temperature. The behaviour of $V_{GS_{th}}$, $V_{GP}$ and $t_{GP}$ in variable temperature and under normal switching remains of great interest, especially with SiC power MOSFETs.

The adaptation of the aforementioned parameters for condition monitoring in the power MOSFET applications has been limited to utilization of the gate-source threshold voltage, stemming from the novelty of identifying $V_{GP}$ and $t_{GP}$ as precursors. In these endeavors the constant current method is typically adopted, where $V_{GS_{th}}$ is achieved at a predetermined drain current, indicating the opening of the channel [108]. During normal operation, this would require an accurate measurement current to be injected at the off-state of the converter, which in turn demands for isolation of the gate from the driver [70]. Considering the high switching frequency of power MOSFETs, which can be even higher with SiC devices, the approach seems difficult and

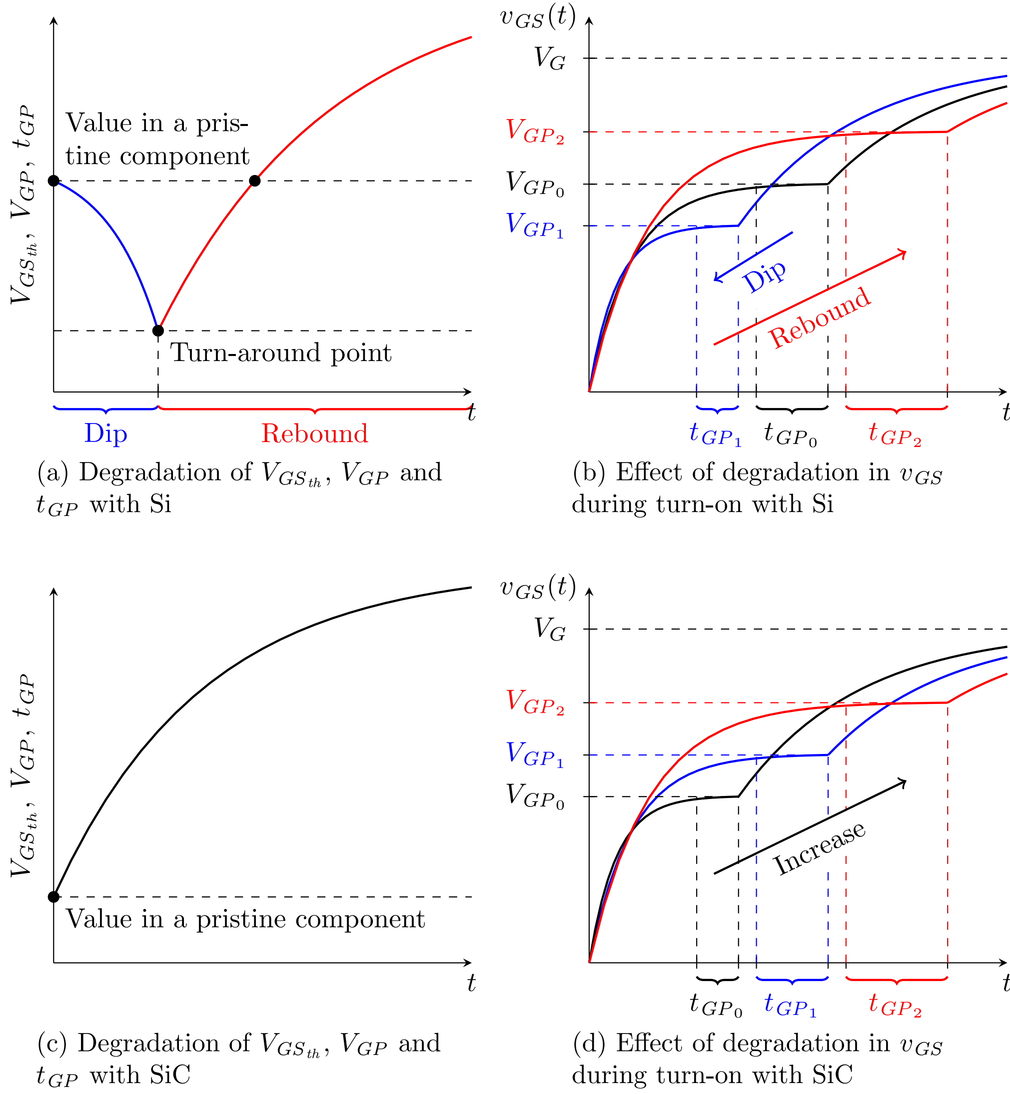



Figure 4.9: Gate oxide degradation precursors in Si (a-b) and SiC (c-d) power MOSFETs.

might be only possible at the times the converter is not modulating [74]. Additionally, the drain current value to be measured in the constant current method is typically very small, which may cause challenges in noisy environments.

# Chapter 5

# Power cycling test

In this chapter we report the power cycling test conducted for accelerated ageing of 11 TO-247-3 packaged SiC power MOSFETs of two different types. Additionally, 3 more devices were saved for an auxiliary cycling session, which is also considered. During the test, the junction temperature of the devices was estimated by the body diode forward voltage for the assessment of the stress profile. The $v_{DS_{on}}$ data collected concurrently with the body diode measurements forms the basis for the prognostic analysis of the on-state resistance documented in Chapter 6.

Since the test system used in this study was just recently procured to the laboratory, its usage was a thorough learning process and setting of some parameters raised particular challenges which we attempted to resolve with careful justification. Proper parameter selection is highly important for achievement of reliable measurements. Furthermore, some parameter values were updated between the test groups by the gained knowledge, in order to generate a more expedient environment for the devices under test (DUT). Therefore, in addition to presenting the test results we review the parameter settings. Exact reporting is also essential for the test repeatability.

## 5.1 Test system

The test equipment used in the study was selected from the already existing apparatus in ABB Drives Quality & Reliability (Q&R) Laboratory. The initial idea was to pursuit for some failure precursor measurements in SiC power MOSFETs during a long-term test. Based on the literature review in the previous chapters, perhaps the most intriguing parameters were related to the Miller platform, considering their close relation with the reliability issues in the gate oxide of SiC devices. However, it was shortly realised that

suitable equipment for such measurements was not readily available.

It was decided to shift the research focus towards packaging-related failure mechanisms as a newly arrived instrument, a customized MicReD® Industrial Power Tester™ 1500A by Mentor® [109], was found suitable in studying some of the related precursors in power MOSFETs. The tester is primarily designed for in-situ failure diagnosis and TSEP-based thermal characterization of various power semiconductors, such as IGBTs, diodes and MOSFETs. The measurements may be further carried out during an automated power cycling test to investigate the signals and the thermal behaviour as the devices are aged. After testing the acquired data can be analysed in multiple manners, the software provided by the manufacturer focuses mostly on deriving and comparing the structure functions [110], which characterize the heat flow through the device. The system supports sensing of several precursors and TSEPs in various loading and powering arrangements.

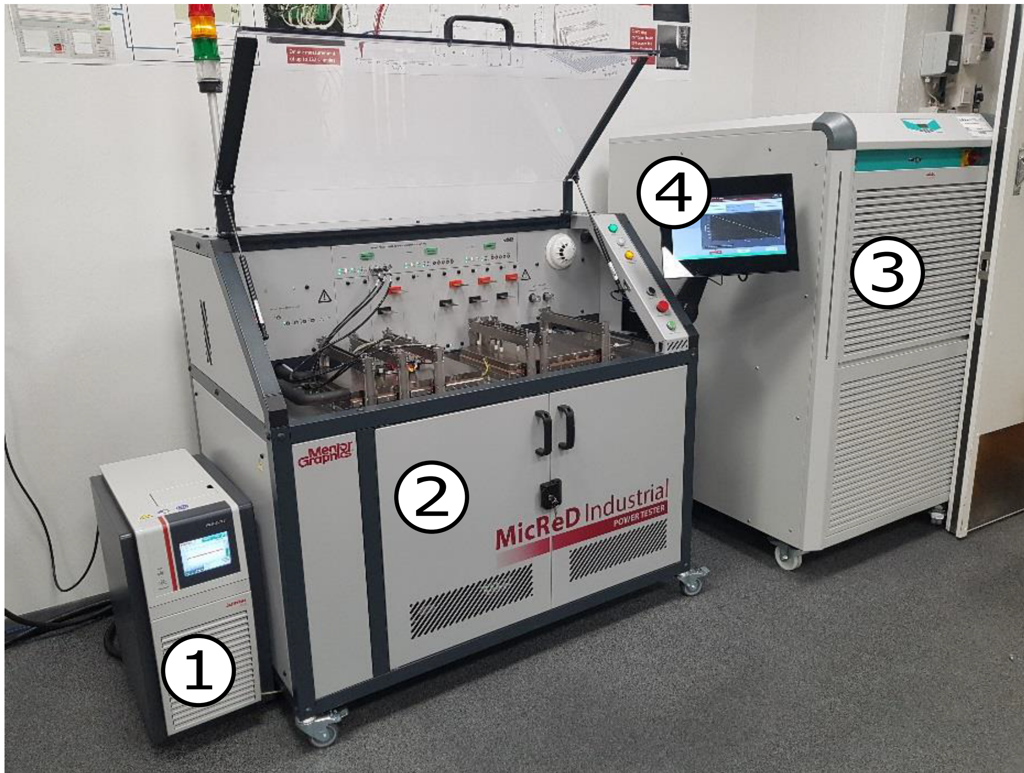



Figure 5.1: The customized MicReD® Industrial Power Tester™ 1500A and the thermostats used in the study.

The test system is shown in total in Figure 5.1, where the actual tester (2) in the middle is accompanied with two heat exchangers used for temperature control of the thermal platforms in the tester. The one on the left (1)

is capable of more accurate and stable regulation and is thus intended for calibration purposes and power cycling of devices requiring only low cooling power, such as the discretes. The one on the right (3), in turn, is to be utilized when power cycling larger semiconductor modules where the amount of the needed cooling power is high. The tester is controlled and configured by the MicRed® Power Tester™ Measurement Control Software which is accessed through the attached touch screen (4).

### 5.1.1 Measurement method

Although in this study we focus on the $R_{DS_{on}}$ as a failure precursor, an additional TSEP measurement is considered to set a suitable temperature profile for the power cycling test. From the several supported configurations in the test system we chose to heat each device under test (DUT) with a high current through the channel and to estimate the junction temperature by the body diode forward voltage measured while the device is cooling. Essentially, the body diode forward voltage $v_F$ is sensed as the drain-source voltage but the measurement result is oppositely signed compared to $v_{DS_{on}}$ due to reversely directed currents.

Figure 5.2: The circuitry for heating the power MOSFET through the channel and sensing the body diode forward voltage.

The wiring diagram of the electrical circuit corresponding the selected measurement method is illustrated in Figure 5.2. The DUT is separated with the dashed contour, whereas the other components are part of the tester hardware. With the help of the figure we can clarify the loading and measuring principles. In the on-state the power MOSFET gate is positively biased with $V_{G_+}$ and a cycling current $I_{cycle}$ is flowed through the channel to heat the component for a user-specified amount of time $t_{on}$. In the off-state the

$I_{cycle}$ is switched off by the tester and the DUT gate is negatively biased with $V_{G_}$, which allows for a constantly applied measurement current $I_{sense}$ to flow through the body diode for the $v_{DS}$ measurement during $t_{off}$. The small measurement current results in much lower power dissipation compared to $I_{cycle}$ and the semiconductor quickly cools down as the package is pressed against the thermal platform kept at low temperature by the heat exchanger. As noticed, the cycling current is controlled by an additional switch, which is operated simultaneously with the gate voltage selector. The measurement current, in turn, is supplied continuously.

The system provides three of the aforementioned measurement channels, whereby three discretely packaged power MOSFETs may be tested independently. The number of the DUTs may be increased, however, as the system provides four measurement positions in each of the channels. In this case, the devices share the same cycling and measurement current, but the gate voltage supplies are individual as well as the voltage measurements, thanks to the multiplexed volt meter. With discretes, this indicates that the devices are to be connected in series. Unfortunately, with the tester used in this study the maximum voltage in each channel was 12 V, which was observed to limit the testing capacity due to high $v_{DS_{on}}$ in SiC power MOSFETs.

To estimate the junction temperature based on the body diode forward voltage, a dependency between the two subjects has to be determined first. This is identified in a calibration procedure where the $I_{sense}$ is constantly applied through the body diode as the DUT is driven into blocking with a user-specified $V_{G_}$. Controlling the thermal platform, the temperature of the power MOSFET is stabilized into a certain value at a time in order to acquire and log the corresponding $v_F$. In the tester the dependency is formed in an automated process by measuring the voltage at multiple user-defined temperatures. The stabilization conditions incorporate monitoring of the thermostat temperature and the $v_F$ measurement, which both must remain within predefined boundaries for a user-specified amount of time. Since the result of the calibration is dependent on the set $I_{sense}$ and $V_{G_}$, it is important to use the same values later on in the power cycling where the dependency is required in converting $v_F$ measurements into temperatures.

During the power cycling the tester measures the body diode forward voltage just after switching the $I_{cycle}$ off and just before switching it back on, as illustrated in Figure 5.3. The logged $v_F$ values, referred to as $V_{hot}$ and $V_{cold}$ respectively, are used in combination with the calibration result to estimate the junction temperatures at the top and the bottom of the cycle, i.e. $T_{j_{max}}$ and $T_{j_{min}}$. Furthermore, for the $T_{j_{max}}$ calculation the electrical transient originating from the switching is compensated with a model fitted in a preliminary transient measurement. The model is extrapolated to the time

instant of the switching to assess the actual $T_{j_{max}}$ at its highest. Finally, the temperature swing $\Delta T_j$ is derived as a difference between the two extremes. At the end of the heating phase the system also measures the $v_{DS_{on}}$, denoted as $V_{on}$, which we used to derive the $R_{DS_{on}}$ data for the prognostic analysis.

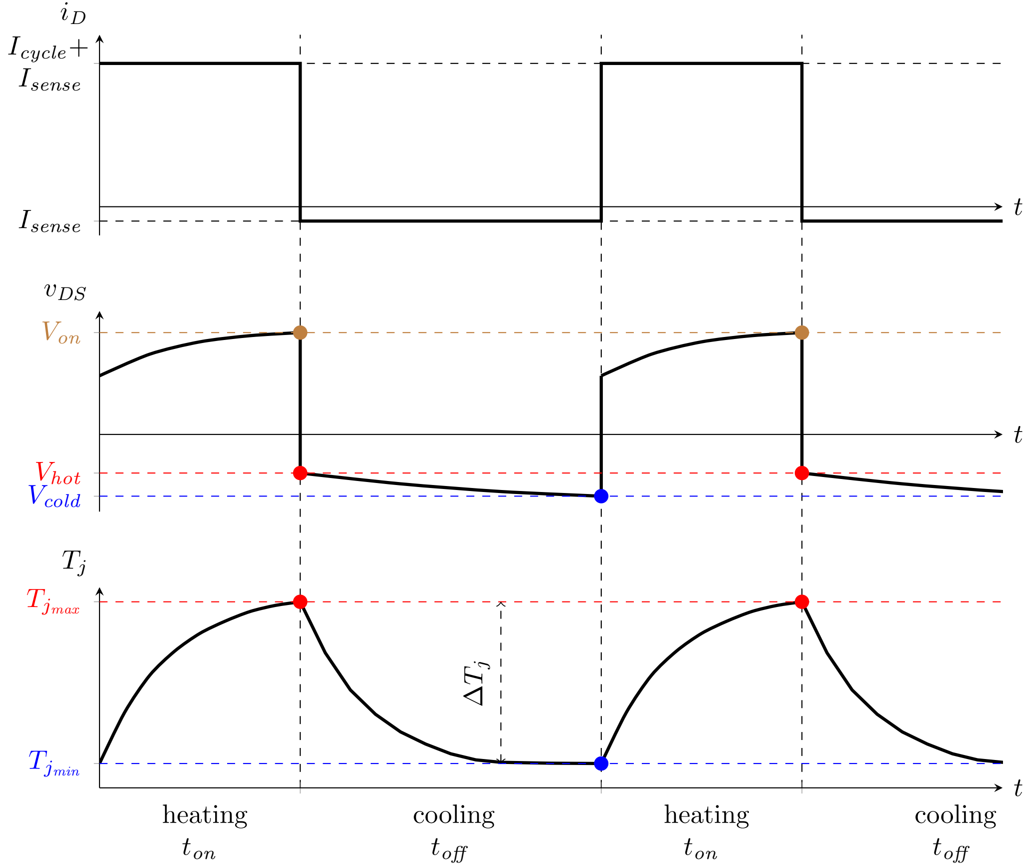



Figure 5.3: The power cycling profile and the measured quantities. In addition, it can be noticed that the actual drain current in the on-state is lower than $I_{cycle}$ as the negative $I_{sense}$ is applied continuously.

Additionally, the user may specify an intermittent $R_{th}$ measurement to characterize the whole cooling curve at desired intervals during the power cycling. Here the test system returns $v_F$ data points during $t_{off}$ with the maximum time resolution of 1 µs [109], thus in much greater amount compared to bare $V_{hot}$ and $V_{cold}$ given in a standard power cycle. The T3ster Master software provided by the manufacturer may be utilized to study the $R_{th}$ measurements and to derive them into the structure functions. In this work this feature was enabled with some of the power cycling sessions, but the collected data was not further analysed.

## 5.2 Test description

Fourteen TO-247-3 packaged SiC power MOSFETs, 9 of the type A and 4 of the type B, were calibrated and power cycled with the introduced test system. In order to set a suitable stress profile, the body diode forward voltage measurement method was utilized to estimate the DUT junction temperatures. To identify each sample we use the notation XXX_Y where XXX is a zero-padded unique ID number for each device, and Y marks the type, A or B in this case.

The 14 samples were divided into three groups. Group 1, i.e. DUTs of IDs 000-006, was tested first and the parameter setting and thus the results reflect the unfamiliarity with the new test system. Group 2, consisting of IDs 007-009, was used to study some of the obscurities arisen with Group 1, and is only briefly considered. For Group 3, comprising the devices of IDs 011-013 and 015, a few new arrangements were made and several parameter values were updated based on the gained understanding from the results of the first two sets. The prognostic analysis documented in Chapter 6 was conducted only for Group 1 and Group 3.

The settings of the test parameters for the DUTs in Group 1 and 3 are collected in Appendix A. In the following subsection 5.2.2 we go through the main points how we came to the presented values. As some of the parameters were updated based on findings in the previous test groups, these changes are considered in the corresponding subsections 5.2.2, 5.2.3 and 5.2.4 also reporting the test results. The cycles-to-failure and selected characteristics of the stress profiles are also summarized Appendix A.

### 5.2.1 Parametrization

Of the user-definable parameters $I_{sense}$ and $V_{G-}$ greatly characterize the resulting calibration curves, and their setting should be considered carefully also in preparation for the power cycling test. For the former we decided to use the maximum negative value available in the tester software, -1000 mA, in order to achieve the strongest possible voltage measurement signal for enhanced noise rejection. A little inconveniently, in the power cycling test this selection was observed to create a trade-off: in the cooling phase the high measurement current heated the DUTs considerably, resulting in elevated $T_{j_{min}}$ and thus lowered $\Delta T_j$. However, by decreasing the temperature of the thermal platform together with increasing the $I_{cycle}$, the problem could be mitigated and desired settings for appropriate temperatures were found. It may be argued, nevertheless, that the greater the measurement current, the

more the cooling behaviour of the DUT is affected, and the more the voltage measurements recede from the *real* values ideally measured without the sensing current. In this study the comparability and quality of the measurement results assumedly provided by the $I_{sense}$ of greater absolute was still preferred.

The $V_{G_}$ was set differently between Groups 1 and 3. For Group 1, the setting was desired to be made equally between the two device types and the value of -3 V was selected for every DUT in that group. However, as some improvements in the behaviour of the body diode forward voltage transient correction with even lower gate-source voltages were noted to have been proposed [111], for Group 3 the voltage was decreased to the lowest operational values allowed in the device data sheets. For the type A this means -5 V and for the type B -6 V.

While setting the parameter values for the power cycling test it is important to keep in mind the basics in accelerated testing. As stated the accelerated conditions should not be traversed too far from the actual use conditions, whereby we have to assess the veracity of the power cycling test to set the associated parameters correctly. Here an important fact to be considered is that as the DUTs are not actively chopping the current, the temperature swings produced are mostly attributed to the conduction losses. Therefore, the power cycling may be interpreted to simulate the start and stop events in a DC-DC converter when the mains are toggled on and off. The test profile is then chiefly comparable to applications with regular starts and stops, with the exception that the smaller temperature fluctuations originating from the normal switching operation are neglected. The life acceleration in the power cycling test is based on more frequently occurring cycles and on the other hand possibly on higher $\Delta T_j$ and $T_{j_{max}}$, depending on the parameter settings. This inference forms the basis for the set parameter values.

The test system supports four different cycling current regulation strategies for the power cycling including constant $I_{cycle}$, constant power step, constant $\Delta T_j$ and constant case temperature swing $\Delta T_c$. Of the available alternatives the first was thought to simulate a real application the best and was therefore selected. This choice, however, led to the question of how to determine the value of $I_{cycle}$, which caused some meaningful difficulties.

The original idea was to set $I_{cycle}$ based on the produced initial $\Delta T_j$, but it was soon observed that even devices of the same type generated very different temperature swings although driven with the same current. This meant that $I_{cycle}$ should have been altered significantly within the DUTs of the same type, which was though to result in abnormal variety of usage considering the regulation in actual applications. Therefore we decided to drive the DUTs of the same type with the same current, but to set its value

so that the average initial $\Delta T_j$ became as desired. Some mitigations to the rule had still to be made in order to not exceed the DUT specifications for the maximum operational $T_j$ to avoid evoking failure mechanisms from unusually high temperatures. To calculate the initial DUT-specific $\Delta T_j$ we used a protocol of ignoring the first 10 measurements to provide a more stable measurement event, and then averaged the subsequent 40 measurements.

Of the remaining power cycling parameters the ones related to the body diode voltage measurement, $I_{sense}$ and $V_{G_-}$, are fixed by the calibration settings. On the contrary, the $V_{G_+}$ has to be determined, and 18 V was selected for every DUT. Considering the DUT specifications, this choice lies in the upper range of the permitted gate-source voltages and was made based on an educated advice.

After setting the parameters, the stop criteria needs to be addressed before starting the power cycling test. For this, the system supports configuring of several limits based on the highest and the lowest permissible measurement values including the maximums in $V_{on}$, $\Delta T_j$ and $T_{j_{max}}$, and the minimums in $V_{on}$, $I_{cycle}$ and $\Delta P$. The limits may be set absolutely, or relatively based on the measurement values at the beginning of the test. In the test system software, these initial values are calculated in default by excluding the first 10 and then averaging the subsequent 10 data points. This procedure was not edited for the study.

Basically, the selection of the limit values depends on the failure definition, which for the most part shares the same challenges as the failure threshold setting already considered for power MOSFET prognostics in 4.2.2. Certain precepts can still be distinguished for research purposes: ideally the stop would occur only when the in-situ measurement data supports the failure event reliably enough, but on the other hand before physical evidence of the failure is destroyed. The former requirement enables the prognostic data analysis whereas the latter is essential for possible after-test root cause investigations of the failed devices. These matters considered, with Group 1 we chose to set the aforementioned maximum limits to 140 % of the initial measurements, whereas the minimum limits were set to 70 % of the initial values. Later on, for Group 3 the maximum limits were updated to 120 % based on findings in the literature [82, 112]. The minimums were considered only to be achieved in the case of problems somewhere else than in the DUTs.

### 5.2.2 Group 1

With Group 1 the targeted initial average $\Delta T_j$ was set to 90 ℃. Considering the DUT specifications, the targeted $\Delta T_j$ represents quite conservative stressing, which was desirable as the power cycling endurance of the samples

was hardly known initially. The rather moderate life acceleration induced by the selected temperature swing was therefore decided to be compensated with a short cycle time to produce cycles at a very high rate. As a consequence, $t_{on}$ and $t_{off}$ were both set to two seconds. The temperature of the thermal platform was kept at 25 °C.

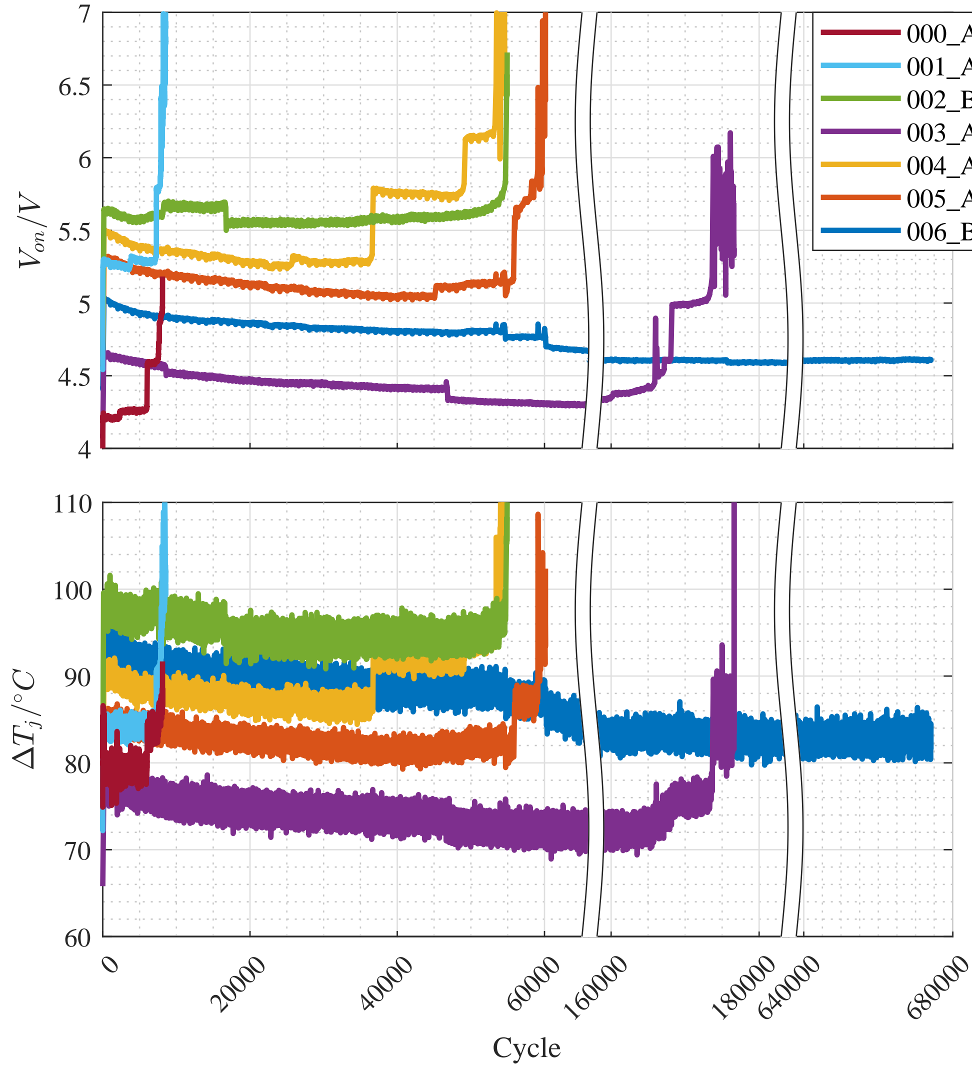



Figure 5.4: Group 1 $V_{on}$ and $\Delta T_j$ measurements. The x-axis has been cut twice to exclude uninteresting sections in the data.

The $V_{on}$ and the $\Delta T_j$ measurements collected during the power cycling test of Group 1 are presented in Figure 5.4, from which the following observations can be made. First, the cycles-to-failure seem to vary significantly between the samples. Second, 006_B did not fail, even though it was run for

over 670000 cycles. Third, it may be noticed that the curves belonging to the type A devices are marked by quite sudden upward jumps after which the curve stays elevated. As discussed in 4.2.2, behaviour of this kind is typically related to the bond wire lift-offs. Also with type B devices some discontinuities exist but these could be traced back to the stops occurred because of failures in other DUTs. Lastly, on the contrary to the expected increasing behaviour, also decrease may be observed in many of the trajectories. This trend seems to decelerate as the test progresses.

The decreasing trend in the $\Delta T_j$ trajectories indicates that the cooling performance of the DUTs enhances during the test. Although this could be considered as a relief in an actual application, from the research perspective it is troublesome as the stress initially set by the targeted $\Delta T_j$ gradually reduces. Furthermore, if some failure mechanisms would induce upward trend, as it was expected, these tendencies might be buried beneath the oppositely directed behaviour, hiding the evidence required in the RUL prediction. For these reasons, we decided to conduct an auxiliary power cycling session for further explanation.

### 5.2.3 Group 2

It was speculated that the decreasing phenomenon could possibly originate from instability in the joint formed by the insulator pad pressed between the power MOSFET back plate and the thermal platform. To study the hypothesis, the Group 2 devices devoted in the test were at first power cycled and then the test was stopped to replace the insulator pads with pristine samples. Afterwards, the test was continued normally. As shown in Figure 5.5, the change of the pads resulted in significant shifts in the $V_{on}$ and $\Delta T_j$ curves towards the initial values at around 35000 cycles. Furthermore, as seen in the figure, the measurement trajectories started to experience similar reduction to the beginning of the test. Obviously, this demonstration does not take into account many other variables in the insulator pad replacement process, but to some extent supports the set hypothesis about the instability in the back-plate-insulator-pad-thermal-platform joint.

A possible solution to the described problem involves a better behaving insulator pad, in terms of thermal stability, which however still requires further research. Another alternative incorporates an additional heat spreader, such as a block of copper, installed between the power MOSFET and the thermal platform. In this case the insulator pad may be moved to the interface of the spreader and the thermal platform, which, as Sarkany et al. [113] state, significantly reduces the effect of the pad. However, it has to be remembered that the latter proposal adds extra variables possibly contributing

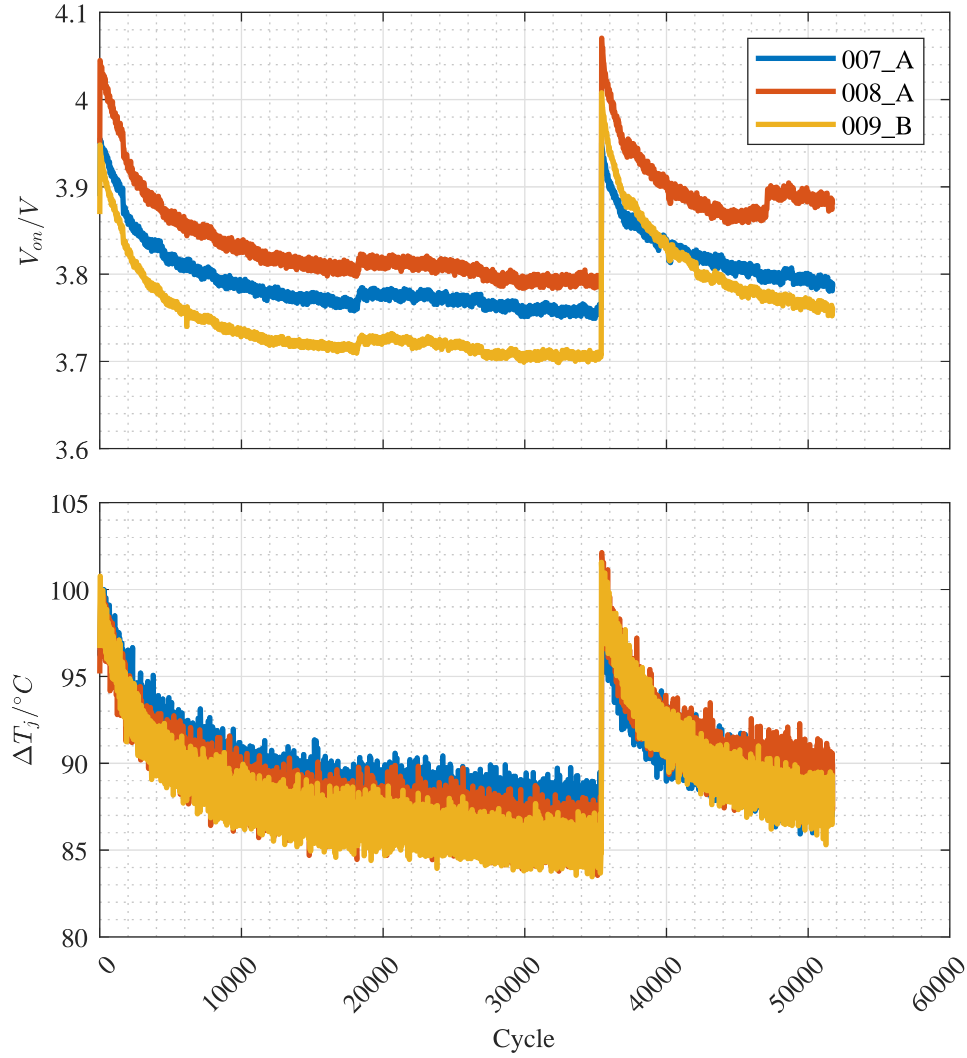



Figure 5.5: Group 2 $V_{on}$ and $\Delta T_j$ measurements. The effect of the replaced insulator pads can be seen at around 35000 cycles.

to the comparability of the results. Due to scope of the work neither of the presented techniques was employed in this study, but they represent possible solutions to the necessary modifications for future investigations.

### 5.2.4 Group 3

As a conclusion of the issues regarded with the power cycling of Group 1 devices, certain new arrangements as well as updates on the parametrization were made for Group 3. Since some of the samples in Group 1 did not fail or

survived very long, the targeted initial $\Delta T_j$ was raised to 125 °C to produce failures more quickly. On the other hand, in order to not exceed the device specific limits, we decided to lower the temperature of the thermal platform to 20 °C. Also the connection method was changed to support more samples simultaneously and two power MOSFETs were now connected in series.

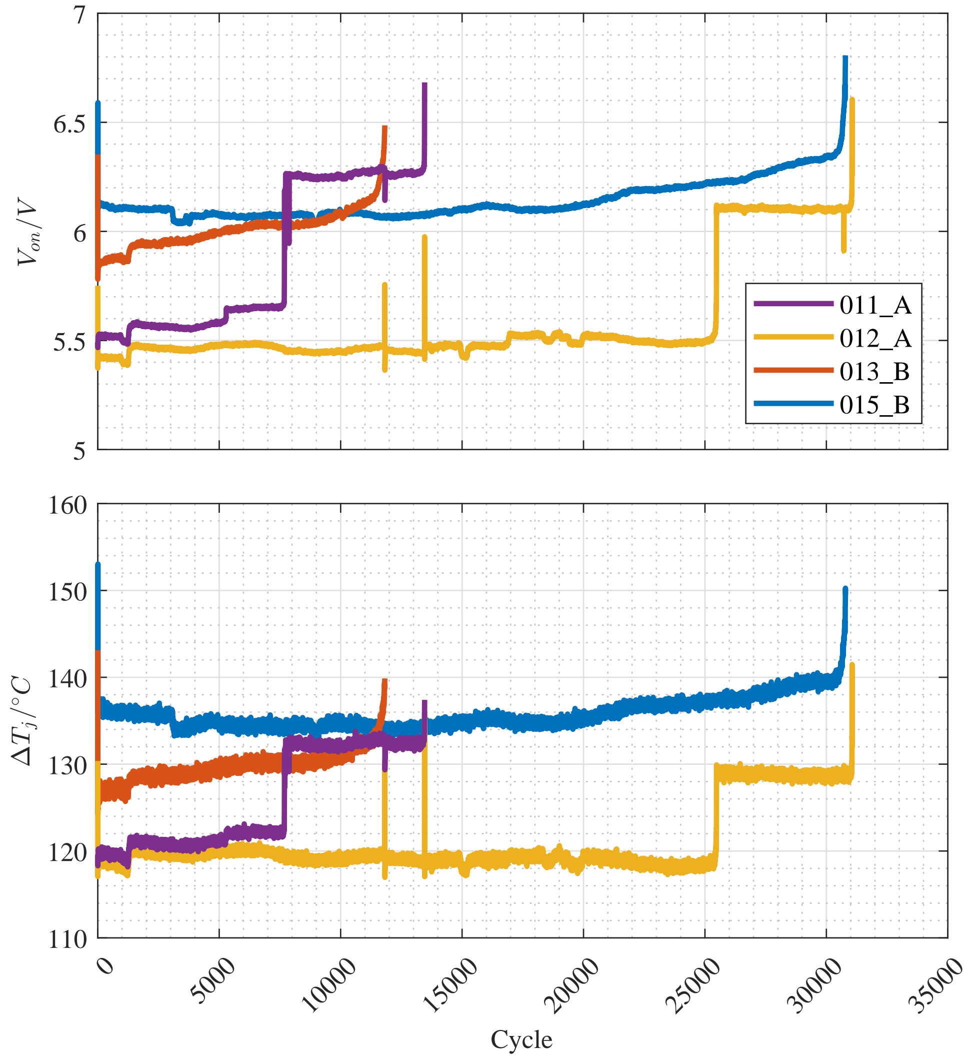



Figure 5.6: Group 3 $V_{on}$ and $\Delta T_j$ measurements. The spikes at the very beginning and at around 12500 cycles were caused by issues with the power compensation in the test system during the start (or resumption) of the test.

It was also be speculated whether the generated stress was overly concentrated on the bond wires in comparison with the actual use conditions of the power MOSFETs. This notion was made as most of the presented trajec-

tories from Group 1 showed resemblance to those in the literature proposed having bond wire issues. Therefore, a decision was made to lengthen the cycle with the aim of relieving the stress experienced by the bond wires and thus possibly target the stress more on the die-attach. As a consequence, $t_{on}$ and $t_{off}$ were set to 5 s and 10 s, respectively.

The $V_{on}$ and the $\Delta T_j$ measurements of Group 3 are shown in Figure 5.6. As noticed, the $\Delta T_j$ is now significantly higher than with Group 1, and all the devices failed within a reasonable time scale. However, although the cycle was lengthened the type A trajectories are still marked by the sudden upturns.

# Chapter 6

# Prognostic analysis

The prognostic analysis was conducted based on the data from the power cycling test of the SiC power MOSFETs in Groups 1 and 3 reported in Chapter 5. More specifically, we derived comparable $\Delta R_{DS_{on}}$ values based on the $V_{on}$ measurements provided by the test system, and developed a particle filtering approach to study physics-based prognostics in the selected DUTs. Furthermore, the suitability of $\Delta R_{DS_{on}}$ as a failure precursor was evaluated in the two tested power MOSFET types. In the chapter below these efforts are documented in detail.

## 6.1 Failure precursor: change in the on-state resistance

To derive $\Delta R_{DS_{on}}$ values from the power cycling data, we utilized the $V_{on}$ measurements provided by the test system. Although in a real application the acquisition of comparable on-state drain-source resistance measurements is challenging, in a controlled environment the task is greatly simplified. With the introduced test system we found three aspects supporting this statement. Firstly, the DUT case temperature is assumed not to vary significantly between the cycles, thanks to the controlled thermal platforms much larger in size compared to the TO-247-3 package. Therefore, the $v_{DS_{on}}$ variation from the external temperature changes may be neglected. Secondly, since the power MOSFETs were driven with a constant cycling current, it is regarded that the total current $I_{cycle} + I_{sense}$ stays similar enough between the $V_{on}$ measurements. Lastly, although the $v_{DS_{on}}$ changes during the on-state due to the increasing temperature, the tester always measures its value, the $V_{on}$, at the end of the on-state.

Combining the above points, we assume that the change in the $V_{on}$ mea-

surements between the cycles is mostly attributed to the ageing effects under investigation and that comparable $R_{DS_{on}}$ values describing those effects may be derived by simply dividing the provided $V_{on}$ with the used current $I_{cycle} + I_{sense}$. Furthermore, to narrow the initial sampling interval of the degradation state set for the estimation algorithm, we can consider the change in $R_{DS_{on}}$ by subtracting each data point with an averaged initial value at the beginning of the trajectory. For the averaging we used the following rule, first 10 samples were neglected and the subsequent 40 were averaged.

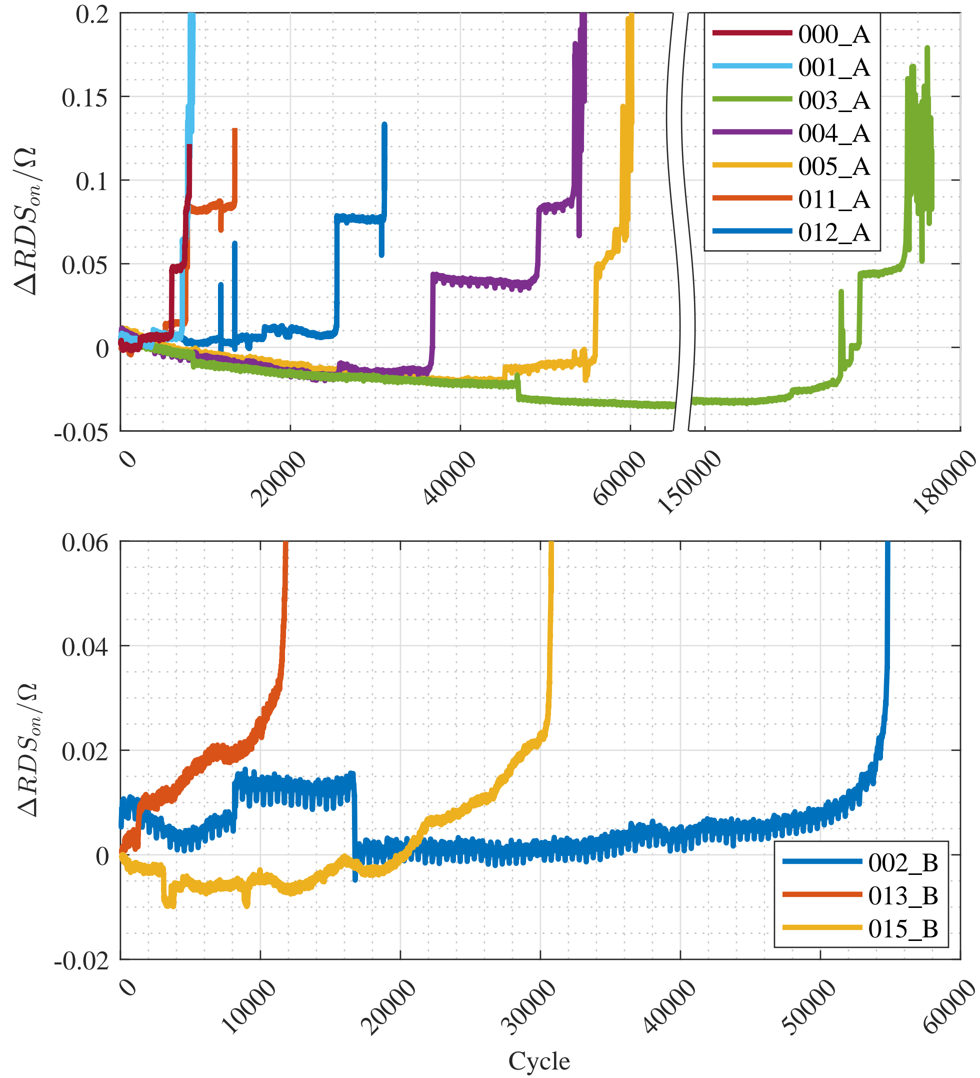



Figure 6.1: The calculated $\Delta R_{DS_{on}}$ of Group 1 and Group 3 divided based on the device type. The sample that survived in the test, 006_B, is excluded.

The $\Delta R_{DS_{on}}$ curves for both Group 1 and 3 are shown in Figure 6.1.

Here we have divide the results in two: the type A devices are shown in the upper graph whereas the lower graph exhibits the curves from type B devices. From the prognostics perspective the former do not seem very potential as the trajectories are marked by the sudden upturns typically observed due to bond-wire lift-offs. Theoretically, as the devices still seem to operate after the first jump, this information could possibly be exploited in simple boolean diagnostics, if the application would only allow such transients. However, the actual RUL prediction by the presented physics-based methods does not seem conceivable considering the lack of observable trend associated with the failure. The type B devices, in turn, experience gradual drift well before the moment of failure, whereby inference of the EoL, and the RUL, in the means of prognostics seems possible. Therefore, in the prognostic analysis we focus only on the type B.

## 6.2 Degradation model development

Due to the lack of actual physical models the $R_{DS_{on}}$ degradation has been typically modelled based on empirically verified behaviour. As introduced in 4.2.2, with traditional Si power MOSFETs it has been observed to follow simple exponential relationships [85, 34, 94]. Since we did not find investigations of $R_{DS_{on}}$ modelling with SiC power MOSFETs in the literature, we decided to adopt the methods proposed with silicon-based devices. Considering the previously presented $\Delta R_{DS_{on}}$ trajectories, this was reagarded feasible as the type B curves seemed to roughly realize exponential behaviour before the actual failure. A modelling example is shown in Figure 6.2, where $\Delta R_{DS_{on}}$ of 013_B is fitted with

$$\Delta R_{DS_{on}}(t) = \theta_1 \exp(\theta_2 t) + \theta_3, \tag{6.1}$$

where $\theta_1$, $\theta_2$ and $\theta_3$ are the model parameters to be determined by the prognostic algorithm. Considering the figure, it can be stated that the model is quite approximate and for example does not take into account the rapid step-like change in the amplitude around 1200 cycles, and neither the change of slope at 6600 cycles. Therefore we have to expect that the parameters of the model are time-variant.

Another possible model, which was also considered, expands (6.1) with another exponential function to characterize the accelerating degradation process close to the moment of failure. The original exponential term, in turn, would be used to describe the gradual change prior the additional term begins to prevail. Similar approach has been applied to prognostics of capacity degradation in Lithium-ion batteries [57]. However, to provide

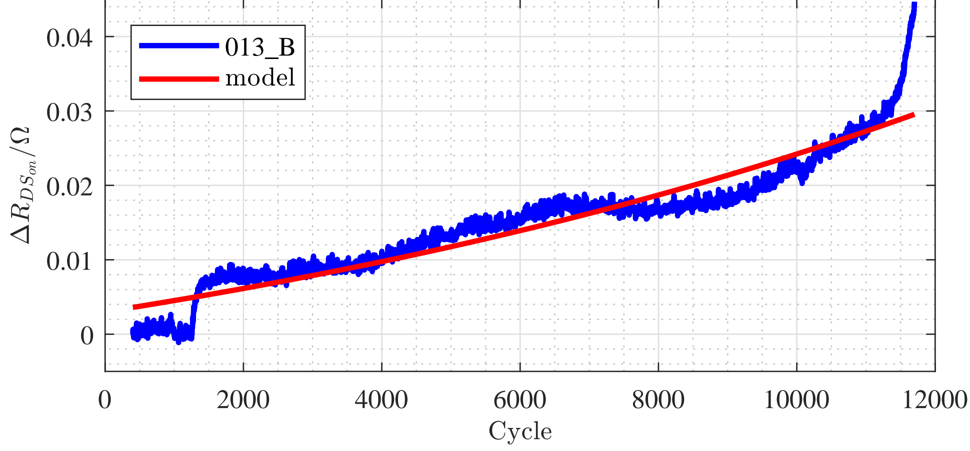



Figure 6.2: $\Delta R_{DS_{on}}$ of 013_B and the model (6.1) with parameters $\theta_1 = 0.02075$, $\theta_2 = 7.041e - 05$ and $\theta_3 = -0.01774$.

some marginal for the prediction uncertainty we speculated that the failure threshold should still situate in the region of safe operation. Consequently, we regarded that it would be appropriate to set the FT at the end part of the gradual degradation, whereby modelling the faster degradation process at the end would not be very meaningful considering the prediction process. Thus, the idea of the double exponential model was omitted and the single exponential model expressed in (6.1) was selected.

## 6.3 Estimation algorithm and metrics

Considering the variety in the presented $\Delta R_{DS_{on}}$ curves and the roughness of the derived model, it seemed essential that the estimation algorithm should be able to track the parameters in parallel with the degradation state. Otherwise successful RUL prediction would require extensive prior knowledge of the parameter values. Regarding the wide variety of environments and usage with actual power semiconductor applications, and also the possibility of the change in the conditions, information this accurate might not be available. Therefore, we adapted a particle filtering approach in pursuit to estimate the joint state-parameter space.

To use the developed degradation model (6.1) with particle filtering, it has to be converted first into discrete state-space. Starting with time-invariant parameters, if we set the state as $x(t) = \exp(\theta_2 t)$ and identify that the measurement $y$ is from $\Delta R_{DS_{on}}$, we can derive the system model in continuous

state-space as

$$\begin{cases} \dot{x}(t) = \theta_2 x(t) + v(t) \\ y(t) = \theta_1 x(t) + \theta_3 + \epsilon(t) . \end{cases} \tag{6.2}$$

Discretization with zero-order hold leads to

$$\begin{cases} x_k = \exp(\theta_2 \Delta t)\, x_{k-1} + v_k \\ y_k = \theta_1 x_k + \theta_3 + \epsilon_k, \end{cases} \tag{6.3}$$

where $\Delta t$ is the time between the steps $k$ and $k-1$.

For the particle filter to estimate the parameters they need to be assigned with some artificial evolution model as discussed in 3.2.2. The typical solution is to use random walk with which particle filter unfortunately results in poor estimation performance due to increasing variance in the estimated posteriors. To address this problem, we applied the ASIR-KS presented in pseudo code for a single iteration in Algorithm 3 in 3.2.2, which exploits kernel smoothing for the parameter posteriors. Shortly, the method shrinks the particles closer to their mean and adds variance controllably based on the smoothing factor $h$. For the resampling procedure we ended up using *multinomial resampling* as it seemed to provide better results in comparison with systematic resampling presented in Algorithm 2 in 3.2.1.

The future degradation is predicted at predetermined points by propagating each particle with the model (6.3) to the failure threshold using the latest state and parameter estimates. An example of the procedure, which was also introduced in 3.3, is presented in Figure 6.3. In the upper graph the actual $\Delta R_{DS_{on}}$ measurements from 013_B are shown in blue, whereas the predicted degradation trajectories are drawn in red. The lower graph, in turn, shows the corresponding EoL distribution given by the predicted FT intersections and the corresponding particle weights evaluated at the prediction point. This discrete probability density is kernel smoothed to provide the presented continuous distribution for enhanced visualization. In the actual analysis we chose to conduct the prediction at 8 points, starting from 20 % of the true EoL and advancing by 10 % percentage point intervals up to 90 %. Furthermore, $\alpha$-$\lambda$ metric was used in analysing the prognostic performance of the developed algorithm. The algorithm and the analysis tools were all programmed with Matlab.

To initialize the estimation algorithm we have to set the distributions from which the first particles are drawn. In the case of typically used uniform distributions a good guess of the distribution limits may enhance the estimation performance significantly. On the other hand, it is often necessary for the algorithm to function satisfactorily although the range would be

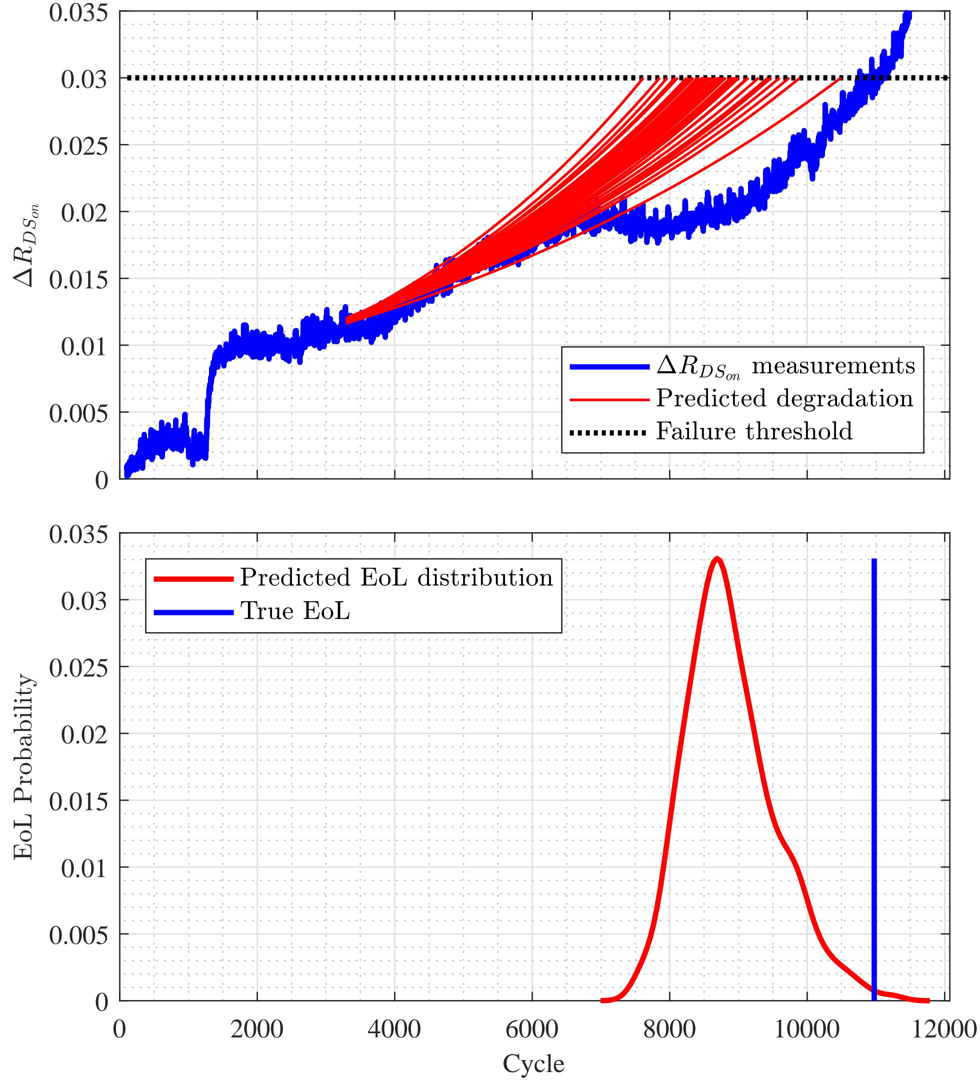



Figure 6.3: An example of the prediction process with 013_B at 30 % of the true EoL. Only 50 of the actual 1000 trajectories are plotted to avoid unnecessarily increasing the file size of the image. The shown EoL distribution is still calculated based on all the particles.

placed with greater uncertainty. In our case, if we assume that the $\Delta R_{DS_{on}}$ increases, then $\theta_1$ and $\theta_2$ would generally be positive. The $\theta_3$, in turn, typically would have a negative value as it may be expected that at the beginning the $\Delta R_{DS_{on}}$ is close to 0. Furthermore, in the case of prognostic algorithms the limits may be specified by assessing the minimum and maximum $t_{EoL}$, and calculating the corresponding parameter values from the degradation model.

## 6.4 Results and discussion

We decided to test our algorithm on the $\Delta R_{DS_{on}}$ measurements collected from the DUT 013_B. This sample was selected as the measurement trajectory incorporated some uncertainties that we considered could be encountered in actual applications as well, namely, the ones discussed in 6.2. On the other hand, this curve did not have the troublesome decreasing phase in the beginning. The change from the decrease to the increase could possible be identified by studying the derivative of the measurement and by starting the estimation only after a possible observation is made. However, in this study we did not want to commit ourselves on handling this issue and thus sample 013_B was selected for the analysis basis. With this trajectory we set the failure threshold in $0.03\Omega$. In possible future investigations the other trajectories may also be considered.

As we observed quite a lot of variance in the RUL prediction performance, we show here results from two distinctive test runs, we call these run 1 and run 2. In these the state and the parameters were initialized as $x = \mathcal{U}[0.9, 1.1]$, $\theta_1 = \mathcal{U}[0, 0.005]$, $\theta_2 = \mathcal{U}[0, 0.0005]$ and $\theta_3 = \mathcal{U}[-0.005, 0]$. Calculating from (6.1) this results in an approximate expected life range of $(3583, \infty)$ cycles. Furthermore, we used 1000 particles, the process noise for $x$ was set to 0.005 and the measurement noise was set to 0.0005. Additionally, smoothing factor of 0.3 was used.

The $\alpha$-$\lambda$ metrics with $\alpha = 0.2$ and $\lambda = 0.5$ for both of the runs are shown in Figures 6.4 and 6.5. Furthermore, the mean, minimum and maximum trajectories of the posterior estimates for the runs 1 and 2 are displayed in Figures 6.6 and 6.7, respectively. The figures are shown in the following three pages.

It seems clear that run 1 performs better in the means of RUL prediction but cannot either stay inside the accuracy cone in the $\alpha$-$\lambda$ metric. In the run 2, in turn, the algorithm underestimates the RUL throughout the run. What is remarkable is that the state estimation performance for the both runs is almost equal although the parameters converge into different values between the runs. Furthermore, the algorithm does not seem to adapt the parameters based to the change in the slope at 6600 cycles. All in all, we were not able to initialize our algorithm in the manner it would have converged into such parameter values that the predicted mean RUL would have stayed inside the $\alpha$-$\lambda$ accuracy zone starting form $t_\lambda$. These points might indicate that the model (6.1) does not describe degradation well enough for the algorithm to decouple such parameter values that would result in successful prediction. On the contrary to the model used here, some physics-based prognostic studies

exploit the actual device physics in the modelling to a much greater extent and map the propagating degradation into the so called wear parameters, which are the parameters to be estimated [46]. However, further research is required before additional conclusions can be made.

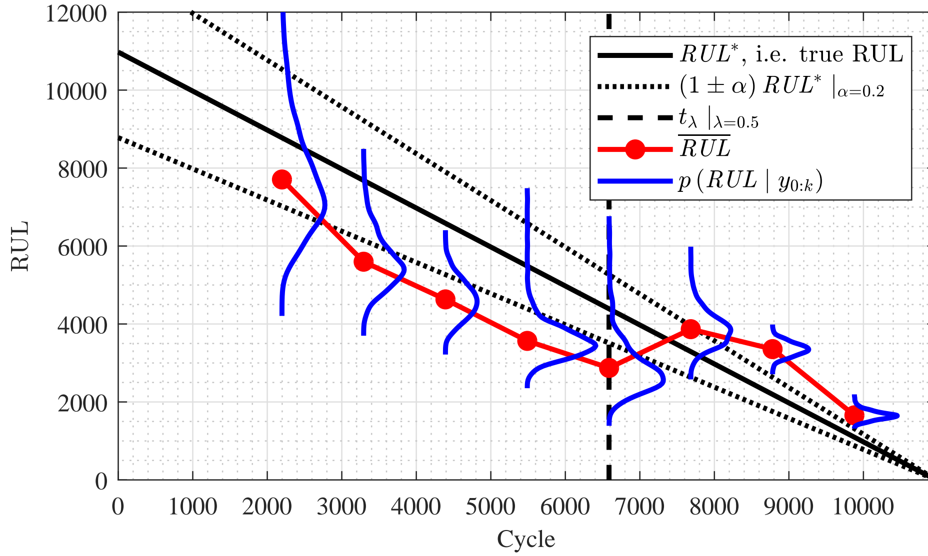



Figure 6.4: Run 1: the $\alpha$-$\lambda$ metric based on the ASIR-KS RUL prediction for $\Delta R_{DS_{on}}$ measurements from 013_B.

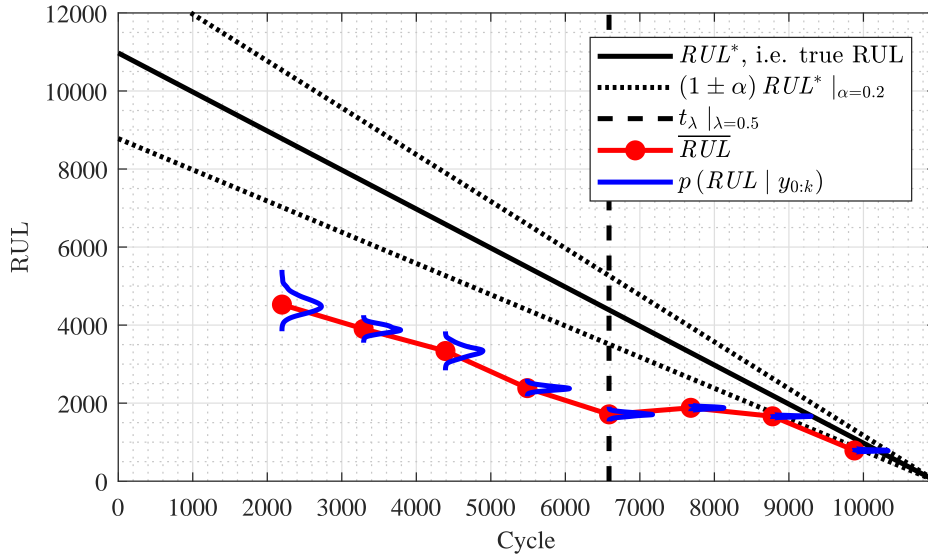



Figure 6.5: Run 2: the $\alpha$-$\lambda$ metric based on the ASIR-KS RUL prediction for $\Delta R_{DS_{on}}$ measurements from 013_B.

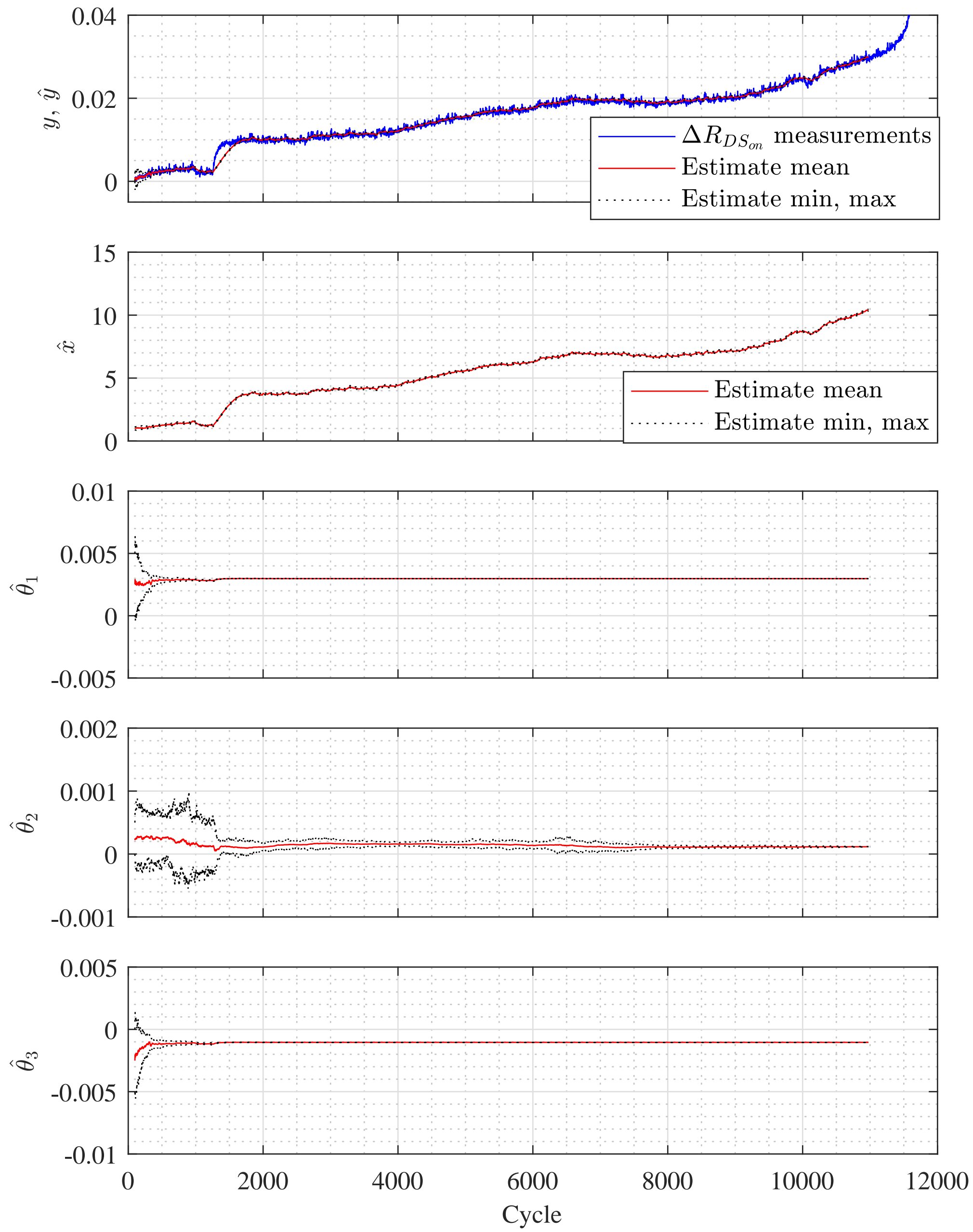



Figure 6.6: Run 1: ASIR-KS posterior estimates based on the $\Delta R_{DS_{on}}$ measurements from 013_B. Estimates from top to bottom: measurement $\hat{y}$, state $\hat{x}$, parameter $\hat{\theta}_1$, parameter $\hat{\theta}_2$ and parameter $\hat{\theta}_3$.

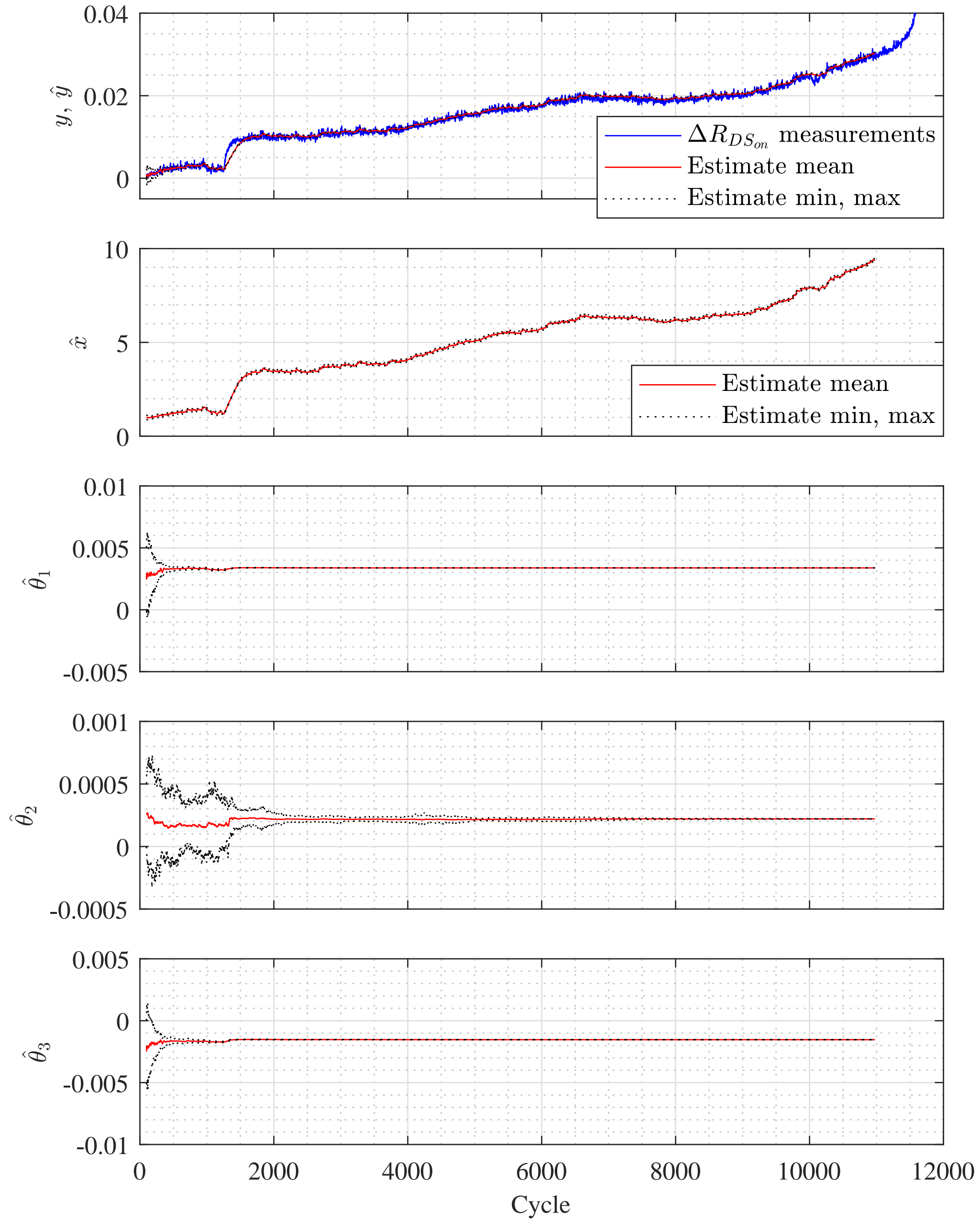



Figure 6.7: Run 2: ASIR-KS posterior estimates based on the $\Delta R_{DS_{on}}$ measurements from 013_B. Estimates from top to bottom: measurement $\hat{y}$, state $\hat{x}$, parameter $\hat{\theta}_1$, parameter $\hat{\theta}_2$ and parameter $\hat{\theta}_3$.

# Chapter 7

# Considerations for the new test system

In order to provide an answer to the final goal set and introduced in Chapter 1, we consider a new test system for reliability and prognostics research in switched-mode DC-DC converters. The SiC power MOSFET, the component studied in this work, acts as a switch in these power electronic applications. The starting point for the considerations is that in the new system the DC-DC converter is tested as a whole.

## 7.1 Testing principles

By the findings in literature and also in the empirical work, we have identified three different principles which the test system may be based on. These principles reflect the possible options we found, how the data analysis within device reliability can generally be conducted. We want to emphasize that the selection of the overall testing principle is one of the most important decisions to be made for the future test design and should not be neglected. Without properly defining the research question, setting the hypothesis and understanding how to analyse the obtained data the development efforts might go to waste.

**Principle 1: Traditional ALT.** The test system may be based on traditional ALT with the emphasis on studying the DC-DC converter as a whole. In this case the time-to-failure data is gathered and analysed as specified in Chapter 2. As the converters may be quite complex, the testing may involve multiple failure mechanisms which are possibly to be separated. The advantages for this system are quite easy implementation, the failure criteria may be set for instance on the power output.

On the other hand, there is always a possibility that no failures occur during the tests, whereupon there would be only little to draw from.

**Principle 2: ALT analysis by critical components.** The second option is to model the reliability of the DC-DC converter by its components. In this case, the converter would be tested as an entity but the time-to-failure models would be based on stress-measurements and failures on individual components identified to be critical for the total device reliability. For instance, we might want to pursuit for TSEP-based temperature measurements in the power MOSFETs.

There are certain considerations, however, that have to be taken into account with this approach. If the stress experienced by a selected critical component did vary within a converter population, due to variability in the device materials, the basics in ALT would require controlling the converters so that similar stress in each selected component is achieved. This is one of the fundamentals in ALT, the analysis is based on statistics where the stress level acts as the common denominator for all the products tested on that level. Again, in the case of the main switch of the converter, if the temperature stress varies within the device population, we could control the electric load to achieve the same temperature in all the main switches. On the other hand, we could also refrain from modifying the electric load and base the analysis on that, in which case the load would be considered the accelerating stress, and the temperature variations would be ignored. Nelson [12, p. 30] describes this selection as identifying the *real stress*, the stress that actually accelerates the failures.

**Principle 3: Studies in prognostics.** The test system may be used for studying the failure precursors for condition-monitoring and prognostics. This option differs from the other two as its purpose is not to build the time-to-failure models but to pursuit for investigations of RUL prediction possibilities. Although the condition monitoring of power electronic components is challenging in the actual use, some complexity may be omitted in the test environment. For instance, the converter could be stopped automatically, the ambient temperature stabilized and the power MOSFET electrically disconnected from the converter to inject a measurement current for more robust $R_{DS_{on}}$ sensing. This approach would generally still need some modifications on the actual application. On the other hand, if the desire is to gather precursor measurements during the operation of the converter the requirements are probably increased. This third principle might be the most difficult

to implement, but on the other hand, might also be the most beneficial considering future investigations.

## 7.2 Selected requirements

As the requirements for the test system are greatly dependent on the chosen testing principle, identified in 7.1, very accurate requirement setting would be inappropriate. Therefore, we present here a guideline which can expanded after the testing principle selection. We have utilized Sonnenfeld et al.'s [114] work among a prognostic test system in specifying some of the requirements.

**Requirement 1: Operation on different types of switched-mode DC-DC converters** The tester must accommodate switched-mode DC-DC converters with different input and output voltages and currents. In order to make the system realizable, some limits for these quantities have to be set. It seems necessary that the system is able to operate at least with every switched-mode DC-DC converter used in the products of ABB Drives, Drive Products.

**Requirement 2: Operation on multiple switched-mode DC-DC converters simultaneously** Considering reliability of the results, the number of the simultaneously tested units has to be high enough in order to provide sufficient amount of data for the analysis. In addition, to shorten the testing time, it is desired that the system is capable of operating on multiple sets of different types of switched-mode DC-DC converters simultaneously.

**Requirement 3: Comprehensive accelerated testing capability** With the test system it must be possible to execute various testing profiles in which elevated temperature and temperature cycling are the main accelerating stresses. In its natural environment, as part of the variable frequency drive, the switched-mode DC-DC converters in question are exposed to temperature-based stresses, stemming from the environment and self-heating from loading.

**Requirement 4: Data acquisition and logging** With the acquisition system it should be possible to reliably gather information about various signals of the device under test and also from the environment. Although the prognostic method should be included in the switched-mode power supply board (in order to enable online prognostics, the measurements in a real implementation have to be done internally),

studying and verification of the measured signals is typically conducted externally.

**Requirement 5: Flexibility and scalability** In a research-oriented environment new insight may often sharpen the actual research question, which may demand reconfiguration of the specified requirements. Thereby, it is best to design the system so that it is easy to modify and to expand. In software terms we can say that the system has to have high cohesion and loose coupling; the system consists of modules, each module's elements have strong relationship to each other, but only need minimum amount of knowledge of the other modules' elements.

The most basic version of the tester consists of a set of modules, which creates a fully functional tester for a certain number of units. In order to cover a greater sample count, it should be able to expand the system by multiplying all or some of these modules. The work needed should only require minor modifications to the the existing components. This applies in both hardware and software. The hardware system should be so flexible that it is easy to add new instruments. In software we are talking about agility: implementation of new software features should not require redesigning of the existing entity.

**Requirement 6: Easy test start-up** To mitigate the effort needed in starting a test, it is desired that the user can set up the test relatively easily. After all, a test system cannot provide any value if it is not used, which might be the case if the system is not easy to set up.

# Chapter 8

# Conclusions

In this study we aimed to demonstrate the RUL prediction process by physics-based prognostic approaches. This was done by developing a kernel-smoothing-based particle filtering algorithm for joint state-parameter estimation and testing it on the drain-source on-state measurement trajectory from the selected power cycled SiC power MOSFET. The results indicated satisfactory performance in the estimation of the states and the parameters but revealed inconsistent performance in the RUL prediction. Although we were not able to produce a solution that would have predicted the RUL in sufficient accuracy in concurrent test cases, the analysis conducted still exemplifies a possible research orientation to be taken.

Furthermore, we wanted to assess the suitability of the developed prognostic method in the two tested device types. The changes in the $\Delta R_{DS_{on}}$ type A devices were so sudden that the prediction by these trajectories seemed very challenging. The type B devices, in turn, showed drift that could possibly be exploited for prognostics. However, also these trajectories showed many uncertainties that have to be considered in further research.

The third objective of the study, the design considerations for a new test system where the power MOSFET would be part of a DC-DC converter, was approached by reviewing general testing principles. The alternatives presented stem from the insight in prognostics and traditional reliability data analysis, gathered throughout the study. Generally, there is a decision to be made whether to base the research on traditional time-to-failure analysis or whether to study the prognostics of power semiconductors devices.

Even though interest in prognostics of power semiconductors has increased it seems that the proposed solutions in the literature are still not very general. An emblematic issue for physics-based approaches is the requirement of a known failure threshold, for which we did not comment on but selected visually a suitable value for the analysis. With actual applica-

tions this is naturally not possible and further research is needed to investigate the matter. Furthermore, the developed method lacked of a well-defined degradation model which may be required for more reliable RUL prediction.

Another group of challenges in the prognostics of power semiconductors stems from the difficulties in the condition monitoring of the precursors, originating from the various dependencies and the connections with multiple failure mechanisms. In this work some of these problems were mitigated by the utilization of the test system but these issues play a significant role in the prognostics of actual applications. The new test system which was given principles for is an example of an approach to study the prognostic solutions in a more natural environment.

# Bibliography


[1] J. W. McPherson. *Reliability Physics and Engineering. Time-to-Failure Modeling*. 2nd ed. Heidelberg: Springer International Publishing, 2013. ISBN: 978-3-319-00122-7. DOI: 10.1007/978-3-319-00122-7.

[2] *Judicious use of aluminum electrolytic capacitors*. Technical Note. Nippon Chemi-Con. URL: https://www.chemi-con.co.jp/e/catalog/pdf/al-e/al-sepa-e/001-guide/al-technote-e-2018.pdf (visited on 30/08/2018).

[3] J. de Vries, M. Jansen and W. van Driel. "On the difference between thermal cycling and thermal shock testing for board level reliability of soldered interconnections". In: *Microelectronics Reliability* 47.2-3 (Feb.–Mar. 2007), pp. 444–449. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2006.05.009.

[4] P. Vassiliou, A. Mettas and T. El-Azzouzi. In: *Handbook of Performability Engineering*. Ed. by K. B. Misra. London: Springer London, 2008. Chap. Quantitative Accelerated Life-testing and Data Analysis, pp. 543–557. ISBN: 978-1-84800-131-2. DOI: 10.1007/978-1-84800-131-2_35.

[5] NIST/SEMATECH. *e-Handbook of Statistical Methods, 8.2.1. How do you choose an appropriate life distribution model?* Oct. 2013. URL: https://www.itl.nist.gov/div898/handbook/apr/section2/apr21.htm (visited on 06/06/2018).

[6] Reliasoft Corporation. *Analysis of Accelerated Life Tests with Competing Failure Modes*. Sept. 2006. URL: http://weibull.com/hotwire/issue67/hottopics67.htm (visited on 06/06/2018).

[7] Reliasoft Corporation. *Life Data Analysis Reference Book. Chapter 4: Parameter Estimation*. Sept. 2015. URL: http://reliawiki.org/index.php/Parameter_Estimation (visited on 09/07/2018).

[8] W. Meeker and L. Escobar. *Statistical Methods for Reliability Data*. Wiley Series in Probability and Statistics. Wiley, 1998. ISBN: 978-0-471-14328-4.

[9] Reliasoft Corporation. *Accelerated Life Testing Data Analysis Reference Book. Appendix B: Parameter Estimation*. Feb. 2017. URL: `http://reliawiki.com/index.php/Appendix_B:_Parameter_Estimation` (visited on 16/06/2018).

[10] N. Balakrishnan and M. Kateri. "On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data". In: *Statistics & Probability Letters* 78.17 (Dec. 2008), pp. 2971–2975. ISSN: 0167-7152. DOI: `10.1016/j.spl.2008.05.019`.

[11] L. P. Gupta, R. C. Gupta and S. J. Lvin. "Numerical methods for the maximum likelihood estimation of weibull parameters". In: *Journal of Statistical Computation and Simulation* 62.1-2 (Dec. 1998), pp. 1–7. DOI: `10.1080/00949659808811921`.

[12] W. Nelson. *Accelerated testing. Statistical models, test plans and data analyses*. John Wiley & Sons, 1990. ISBN: 0-471-52277-5.

[13] L. A. Escobar and W. Q. Meeker. "A review of accelerated test models". In: *Statistical Science* 21.4 (Nov. 2006), pp. 552–577. ISSN: 0883-4237. DOI: `10.1214/088342306000000321`.

[14] S. Yang, D. Xiang, A. Bryant, P. Mawby, L. Ran and P. Tavner. "Condition Monitoring for Device Reliability in Power Electronic Converters: A Review". In: *IEEE Transactions on Power Electronics* 25.11 (Nov. 2010), pp. 2734–2752. ISSN: 0885-8993. DOI: `10.1109/TPEL.2010.2049377`.

[15] N. M. Vichare and M. G. Pecht. "Prognostics and health management of electronics". In: *IEEE Transactions on Components and Packaging Technologies* 29.1 (Mar. 2006), pp. 222–229. ISSN: 1521-3331. DOI: `10.1109/TCAPT.2006.870387`.

[16] N.-H. Kim, D. An and J.-H. Choi. *Prognostics and Health Management of Engineering Systems. An Introduction*. Springer International Publishing, 2017. ISBN: 978-3-319-44742-1. DOI: `10.1007/978-3-319-44742-1`.

[17] K. Javed, R. Gouriveau and N. Zerhouni. "State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels". In: *Mechanical Systems and Signal Processing* 94 (Sept. 2017), pp. 214–236. ISSN: 0888-3270. DOI: `10.1016/j.ymssp.2017.01.050`.

[18] K. Goebel, B. Saha, A. Saxena, J. R. Celaya and J. P. Christophersen. “Prognostics in Battery Health Management”. In: *IEEE Instrumentation Measurement Magazine* 11.4 (Aug. 2008), pp. 33–40. ISSN: 1094-6969. DOI: 10.1109/MIM.2008.4579269.

[19] S. M. Rezvanizaniani, Z. Liu, Y. Chen and J. Lee. “Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility”. In: *Journal of Power Sources* 256 (June 2014), pp. 110–124. ISSN: 0378-7753. DOI: 10.1016/j.jpowsour.2014.01.085.

[20] M. Berecibar, I. Gandiaga, I. Villarreal, N. Omar, J. V. Mierlo and P. V. den Bossche. “Critical review of state of health estimation methods of Li-ion batteries for real applications”. In: *Renewable and Sustainable Energy Reviews* 56 (Apr. 2016), pp. 572–587. ISSN: 1364-0321. DOI: 10.1016/j.rser.2015.11.042.

[21] L. Wu, J. Yang, Z. Peng and H. Wang. “Remaining useful life prognostic of power metal oxide semiconductor field effect transistor based on improved particle filter algorithm”. In: *Advances in Mechanical Engineering* 9.12 (Dec. 2017). DOI: 10.1177/1687814017749324.

[22] J. R. Celaya, A. Saxena, S. Saha, V. Vashchenko and K. Goebel. “Prognostics of power MOSFET”. In: *Proceedings of the 23rd International Symposium on Power Semiconductor Devices and ICs (ISPSD)* (San Diego, CA, USA, 23–26 May 2011). 2011, pp. 160–163. ISBN: 978-1-4244-8424-9. DOI: 10.1109/ISPSD.2011.5890815.

[23] N. Degrenne, J. Ewanchuk, E. David, R. Boldyrjew and S. Mollov. “A review of prognostics and health management for power semiconductor modules”. In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2015 (PHM)* (New Orleans, LA, USA, 19–24 Oct. 2015). 2015, pp. 242–252. ISBN: 978-1-936263-20-2. URL: https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2015/phmc_15_067.pdf (visited on 24/08/2018).

[24] C. S. Kulkarni, J. R. Celaya, G. Biswas and K. Goebel. “Accelerated aging experiments for capacitor health monitoring and prognostics”. In: *Proceedings of the 2012 IEEE AUTOTESTCON* (Anaheim, CA, USA, 10–13 Sept. 2012). 2012, pp. 356–361. ISBN: 978-1-4673-0700-0. DOI: 10.1109/AUTEST.2012.6334580.

[25] N. Khera and S. A. Khan. “Prognostics of aluminum electrolytic capacitors using artificial neural network approach”. In: *Microelectronics Reliability* 81 (Feb. 2018), pp. 328–336. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2017.11.002.

[26] M. Rigamonti, P. Baraldi, E. Zio, D. Astigarraga and A. Galarza. "Particle Filter-Based Prognostics for an Electrolytic Capacitor Working in Variable Operating Conditions". In: *IEEE Transactions on Power Electronics* 31.2 (Feb. 2016), pp. 1567–1575. ISSN: 0885-8993. DOI: 10.1109/TPEL.2015.2418198.

[27] T. Sutharssan, C. Bailey, S. Stoyanov and Y. Rosunally. "Prognostics and reliability assessment of light emitting diode packaging". In: *Proceedings of the 12th International Conference on Electronic Packaging Technology and High Density Packaging (ICEPT)* (Shanghai, China, 8–11 Aug. 2011). 2011, pp. 1–7. ISBN: 978-1-4577-1768-0. DOI: 10.1109/ICEPT.2011.6066984.

[28] M. H. Chang, M. Kang and M. Pecht. "Prognostics-Based LED Qualification Using Similarity-Based Statistical Measure With RVM Regression Model". In: *IEEE Transactions on Industrial Electronics* 64.7 (July 2017), pp. 5667–5677. ISSN: 0278-0046. DOI: 10.1109/TIE.2017.2677301.

[29] J. Fan, K. C. Yung and M. Pecht. "Physics-of-Failure-Based Prognostics and Health Management for High-Power White Light-Emitting Diode Lighting". In: *IEEE Transactions on Device and Materials Reliability* 11.3 (Sept. 2011), pp. 407–416. ISSN: 1530-4388. DOI: 10.1109/TDMR.2011.2157695.

[30] D. An, N. H. Kim and J.-H. Choi. "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews". In: *Reliability Engineering & System Safety* 133 (Jan. 2015), pp. 223–236. ISSN: 0951-8320. DOI: 10.1016/j.ress.2014.09.014.

[31] J. Sikorska, M. Hodkiewicz and L. Ma. "Prognostic modelling options for remaining useful life estimation by industry". In: *Mechanical Systems and Signal Processing* 25.5 (July 2011), pp. 1803–1836. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2010.11.018.

[32] S. Dusmez, H. Duran and B. Akin. "Remaining Useful Lifetime Estimation for Thermally Stressed Power MOSFETs Based on on-State Resistance Variation". In: *IEEE Transactions on Industry Applications* 52.3 (May 2016), pp. 2554–2563. ISSN: 0093-9994. DOI: 10.1109/TIA.2016.2518127.

[33] P. Lall, J. Wei and K. Goebel. "Comparison of Lalman-filter and extended Kalman-filter for prognostics health management of electronics". In: *Proceedings of the 13th InterSociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*

(San Diego, CA, USA, 30 May–1 June 2012). 2012, pp. 1281–1291. ISBN: 978-1-4244-9531-3. DOI: 10.1109/ITHERM.2012.6231569.

[34] J. Celaya, A. Saxena, S. Saha and K. Goebel. "Prognostics of power mosfets under thermal stress accelerated aging using data-driven and model-based methodologies". In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2011 (PHM)* (Montreal, Canada, 25–29 Sept. 2011). 2011, pp. 443–452. ISBN: 978-1-936263-03-5. URL: https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2011/phmc_11_009.pdf (visited on 24/08/2018).

[35] C. Kulkarni, J. Celaya, K. Goebel and G. Biswas. "Bayesian framework approach for prognostic studies in electrolytic capacitor under thermal overstress conditions". In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2012 (PHM)* (Minneapolis, MN, USA, 23–27 Sept. 2012). 2012, pp. 152–162. ISBN: 978-1-936263-05-9. URL: https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2012/phmc_12_117.pdf (visited on 24/08/2018).

[36] J. Choi, D. An, J. Gang, J. Joo and N. Kim. "Bayesian approach for parameter estimation in the structural analysis and prognosis". In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2010 (PHM)* (Portland, OR, USA, 13–16 Oct. 2010). 2010. ISBN: 978-1-936263-01-1. URL: https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2010/phmc_10_084.pdf (visited on 24/08/2018).

[37] T. Bayes and R. Price. "An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F. R. S. communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S". In: *Philosophical Transactions of the Royal Society of London* 53 (1763), pp. 370–418. DOI: 10.1098/rstl.1763.0053.

[38] A. Doucet, N. de Freitas and N. Gordon. "An Introduction to Sequential Monte Carlo Methods". In: *Sequential Monte Carlo Methods in Practice*. Ed. by A. Doucet, N. de Freitas and N. Gordon. New York, NY: Springer New York, 2001, pp. 3–14.

[39] B. Ristic, S. Arulampalam and N. Gordon. *Beyond the Kalman filter. Particle filters for tracking applications*. Artech House, 2004. ISBN: 1-58053-631-x.

[40] N. J. Gordon, D. J. Salmond and A. F. M. Smith. "Novel approach to nonlinear/non-Gaussian Bayesian state estimation". In: *IEE Proceedings F - Radar and Signal Processing* 140.2 (Apr. 1993), pp. 107–113. ISSN: 0956-375X. DOI: 10.1049/ip-f-2.1993.0015.

[41] L. Olivier, B. Huang and I. Craig. "Dual particle filters for state and parameter estimation with application to a run-of-mine ore mill". In: *Journal of Process Control* 22.4 (Apr. 2012), pp. 710–717. ISSN: 0959-1524. DOI: 10.1016/j.jprocont.2012.02.009.

[42] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra and N. Zerhouni. "Particle filter-based prognostics: Review, discussion and perspectives". In: *Mechanical Systems and Signal Processing* 72-73 (May 2016), pp. 2–31. ISSN: 0888-3270. DOI: 10.1016/j.ymssp.2015.11.008.

[43] R. Douc and O. Cappe. "Comparison of resampling schemes for particle filtering". In: *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis (ISPA)* (Zagreb, Croatia, 15–17 Sept. 2005). 2005, pp. 64–69. ISBN: 953-184-089-X. DOI: 10.1109/ISPA.2005.195385.

[44] J. Carpenter, P. Clifford and P. Fearnhead. "Improved particle filter for nonlinear problems". English. In: *IEE Proceedings - Radar, Sonar and Navigation* 146.1 (Feb. 1999), 2–7(5). ISSN: 1350-2395. DOI: 10.1049/ip-rsn:19990255.

[45] D. Whitley. "A genetic algorithm tutorial". In: *Statistics and Computing* 4.2 (June 1994), pp. 65–85. ISSN: 1573-1375. DOI: 10.1007/BF00175354.

[46] M. J. Daigle and K. Goebel. "Model-Based Prognostics With Concurrent Damage Progression Processes". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 43.3 (May 2013), pp. 535–546. ISSN: 2168-2216. DOI: 10.1109/TSMCA.2012.2207109.

[47] G. Kitagawa. "A Self-Organizing State-Space Model". In: *Journal of the American Statistical Association* 93.443 (Sept. 1998), pp. 1203–1215. ISSN: 01621459. DOI: 10.2307/2669862.

[48] P. Li, R. Goodall and V. Kadirkamanathan. "Parameter estimation of railway vehicle dynamic model using rao-blackwellised particle filter". In: *Proceedings of the 2003 European Control Conference (ECC)* (Cambridge, UK, 1–4 Sept. 2003). 2003, pp. 2384–2389. ISBN: 978-3-9524173-7-9. DOI: 10.23919/ECC.2003.7085323.

[49] D. An, J.-H. Choi and N. Kim. "A comparison study of methods for parameter estimation in the physics-based prognostics". In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2012 (PHM)* (Minneapolis, MN, USA, 23–27 Sept. 2012). 2012. ISBN: 978-1-936263-05-9. URL: `http://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2012/phmc_12_120.pdf` (visited on 24/08/2018).

[50] A. Doucet and V. B. Tadić. "Parameter estimation in general state-space models using particle methods". In: *Annals of the Institute of Statistical Mathematics* 55.2 (June 2003), pp. 409–422. ISSN: 1572-9052. DOI: `10.1007/BF02530508`.

[51] J. Liu and M. West. "Combined Parameter and State Estimation in Simulation-Based Filtering". In: *Sequential Monte Carlo Methods in Practice*. Ed. by A. Doucet, N. de Freitas and N. Gordon. New York, NY: Springer New York, 2001, pp. 197–223.

[52] C. Cheng and J. Tourneret. "A kernel density-based particle filter for state and time-varying parameter estimation in nonlinear state-space models". In: *Proceedings of the 25th European Signal Processing Conference (EUSIPCO)* (Kos, Greece, 28 Aug.–2 Sept. 2017). 2017, pp. 1664–1668. ISBN: 978-0-9928626-7-1. DOI: `10.23919/EUSIPCO.2017.8081492`.

[53] W. Wang, S. Liao and T. Xing. "Particle filter for state and parameter estimation in passive ranging". In: *2009 IEEE International Conference on Intelligent Computing and Intelligent Systems*. Vol. 3. Nov. 2009, pp. 257–261. DOI: `10.1109/ICICISYS.2009.5358175`.

[54] M. K. Pitt and N. Shephard. "Filtering via Simulation: Auxiliary Particle Filters". In: *Journal of the American Statistical Association* 94.446 (1999), pp. 590–599. DOI: `10.1080/01621459.1999.10474153`.

[55] T. Chen, J. Morris and E. Martin. "Particle filters for state and parameter estimation in batch processes". In: *Journal of Process Control* 15.6 (Sept. 2005), pp. 665–673. ISSN: 0959-1524. DOI: `10.1016/j.jprocont.2005.01.001`.

[56] A. Tulsyan, B. Huang, R. B. Gopaluni and J. F. Forbes. "On simultaneous on-line state and parameter estimation in non-linear state-space models". In: *Journal of Process Control* 23.4 (Apr. 2013), pp. 516–526. ISSN: 0959-1524. DOI: `10.1016/j.jprocont.2013.01.010`.

[57] Y. Hu, P. Baraldi, F. D. Maio and E. Zio. “A particle filtering and kernel smoothing-based approach for new design component prognostics”. In: *Reliability Engineering & System Safety* 134 (Feb. 2015), pp. 19–31. ISSN: 0951-8320. DOI: 10.1016/j.ress.2014.10.003.

[58] X. Liu, Z. Chen, C. Zhang and J. Wu. “A novel temperature-compensated model for power Li-ion batteries with dual-particle-filter state of charge estimation”. In: *Applied Energy* 123 (June 2014), pp. 263–272. ISSN: 0306-2619. DOI: 10.1016/j.apenergy.2014.02.072.

[59] A. Saxena, J. Celaya, B. Saha, S. Saha and K. Goebel. “On applying the prognostic performance metrics”. In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2009 (PHM)* (San Diego, CA, USA, 27 Sept.–1 Oct. 2009). 2009. ISBN: 978-1-936263-00-4. URL: https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2009/phmc_09_39.pdf (visited on 24/08/2018).

[60] J. Lutz, H. Schlangenotto, U. Scheuermann and R. De Doncker. *Semiconductor Power Devices: Physics, Characteristics, Reliability.* Springer Berlin Heidelberg, 2011, pp. 283–314. ISBN: 978-3-642-11125-9. DOI: 10.1007/978-3-642-11125-9.

[61] B. J. Baliga. *Fundamentals of Power Semiconductor Devices.* Springer Science & Business Media, 2008. ISBN: 978-0-387-47314-7.

[62] D. Prasad, D. Kastha, S. Sengupta and N. De. *Module 1: Power Semiconductor Devices. Lesson 6: Metal Oxide Semiconductor Field Effect Transistor (MOSFET).* Lecture Notes. 2009. URL: https://nptel.ac.in/courses/108105066/6 (visited on 26/08/2018).

[63] R. Perret. *Power Electronics Semiconductor Devices.* Wiley, 2009. ISBN: 978-1-848-21064-6.

[64] S. Havanur. *Power MOSFET Basics: Understanding the Turn-On Process.* Tech. rep. AN850. Vishay Siliconix, June 2015. URL: https://www.vishay.com/docs/68214/turnonprocess.pdf (visited on 24/08/2018).

[65] D. Cittanti, F. Iannuzzo, E. Hoene and K. Klein. “Role of parasitic capacitances in power MOSFET turn-on switching speed limits: A SiC case study”. In: *Proceedings of the 2017 IEEE Energy Conversion Congress and Exposition (ECCE)* (Cincinnati, OH, USA, 1–5 Oct. 2017). 2017, pp. 1387–1394. ISBN: 978-1-5090-2998-3. DOI: 10.1109/ECCE.2017.8095952.

[66] V. Barkhordarian. *Power MOSFET Basics*. Tech. rep. AN-1084. Infineon. URL: https://www.infineon.com/dgdl/an-1084.pdf?fileId=5546d462533600a401535595840d101e (visited on 24/08/2018).

[67] Z. Chen. "Characterization and modeling of high-switching-speed behavior of SiC active devices". MA thesis. Virginia Tech, 2009. URL: https://vtechworks.lib.vt.edu/bitstream/handle/10919/30778/Chen_Z_T_2009.pdf (visited on 15/03/2018).

[68] H. Li, X. Liao, Y. Hu, Z. Huang and K. Wang. "Analysis of Voltage Variation in Silicon Carbide MOSFETs during Turn-On and Turn-Off". In: *Energies* 10.10 (Sept. 2017). ISSN: 1996-1073. DOI: 10.3390/en10101456.

[69] G. L. Skibinski and W. A. Sethares. "Thermal parameter estimation using recursive identification". In: *IEEE Transactions on Power Electronics* 6.2 (Apr. 1991), pp. 228–239. ISSN: 0885-8993. DOI: 10.1109/63.76809.

[70] H. Chen, B. Ji, V. Pickert and W. Cao. "Real-Time Temperature Estimation for Power MOSFETs Considering Thermal Aging Effects". In: *IEEE Transactions on Device and Materials Reliability* 14.1 (Mar. 2014), pp. 220–228. ISSN: 1530-4388. DOI: 10.1109/TDMR.2013.2292547.

[71] D. L. Blackburn. "Temperature measurements of semiconductor devices - a review". In: *Twentieth Annual IEEE Semiconductor Thermal Measurement and Management Symposium (IEEE Cat. No.04CH37545)*. Mar. 2004, pp. 70–80. DOI: 10.1109/STHERM.2004.1291304.

[72] Y. Avenas, L. Dupont and Z. Khatir. "Temperature Measurement of Power Semiconductor Devices by Thermo-Sensitive Electrical Parameters - A Review". In: *IEEE Transactions on Power Electronics* 27.6 (June 2012), pp. 3081–3092. ISSN: 0885-8993. DOI: 10.1109/TPEL.2011.2178433.

[73] M. Ciappa. "Selected failure mechanisms of modern power modules". In: *Microelectronics Reliability* 42.4-5 (Apr.–May 2002), pp. 653–667. ISSN: 0026-2714. DOI: 10.1016/S0026-2714(02)00042-2.

[74] Y. Avenas, L. Dupont, N. Baker, H. Zara and F. Barruel. "Condition Monitoring: A Decade of Proposed Techniques". In: *IEEE Industrial Electronics Magazine* 9.4 (Dec. 2015), pp. 22–36. ISSN: 1932-4529. DOI: 10.1109/MIE.2015.2481564.

[75] S. Dusmez and B. Akin. “An Active Life Extension Strategy for Thermally Aged Power Switches Based on the Pulse-Width Adjustment Method in Interleaved Converters”. In: *IEEE Transactions on Power Electronics* 31.7 (July 2016), pp. 5149–5160. ISSN: 0885-8993. DOI: 10.1109/TPEL.2015.2487263.

[76] Y. Xiong, X. Cheng, Z. J. Shen, C. Mi, H. Wu and V. K. Garg. “Prognostic and Warning System for Power-Electronic Modules in Electric, Hybrid Electric, and Fuel-Cell Vehicles”. In: *IEEE Transactions on Industrial Electronics* 55.6 (June 2008), pp. 2268–2276. ISSN: 0278-0046. DOI: 10.1109/TIE.2008.918399.

[77] M. Mattila. “Lämpötilan vaihteluiden vaikututkset vaihtosuuntaajan IGBT-moduuliin (Insulated Gate Bipolar Transistor)”. MA thesis. Teknillinen Korkeakoulu, 2008. URL: http://lib.tkk.fi/Dipl/2008/urn012395.pdf (visited on 24/08/2018).

[78] E. Ugur and B. Akin. “Aging assessment of discrete SiC MOSFETs under high temperature cycling tests”. In: *Proceedings of the 2017 IEEE Energy Conversion Congress and Exposition (ECCE)* (Cincinnati, OH, USA, 1–5 Oct. 2017). 2017, pp. 3496–3501. ISBN: 978-1-5090-2998-3. DOI: 10.1109/ECCE.2017.8096624.

[79] V. Sarihan. “Energy based methodology for damage and life prediction of solder joints under thermal cycling”. In: *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B* 17.4 (Nov. 1994), pp. 626–631. ISSN: 1070-9894. DOI: 10.1109/96.338733.

[80] D. C. Katsis and J. D. vanWyk. “A thermal, mechanical, and electrical study of voiding in the solder die-attach of power MOSFETs”. In: *IEEE Transactions on Components and Packaging Technologies* 29.1 (Mar. 2006), pp. 127–136. ISSN: 1521-3331. DOI: 10.1109/TCAPT.2005.853301.

[81] A. Sow, S. Somaya, Y. Ousten, J.-M. Vinassa and F. Patoureaux. “Power MOSFET active power cycling for medical system reliability assessment”. In: *Microelectronics Reliability* 53.9-11 (Sept.–Nov. 2013). European Symposium on Reliability of Electron Devices, Failure Physics and Analysis, pp. 1697–1702. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2013.07.096.

[82] R. Schmidt and U. Scheuermann. “Separating Failure Modes in Power Cycling Tests”. In: *Proceedings of the 7th International Conference on Integrated Power Electronics Systems (CIPS)* (Nuremberg, Germany, 6–8 Mar. 2012). 2012, pp. 1–6. ISBN: 978-3-8007-3414-6. URL: https://ieeexplore.ieee.org/document/6170628/ (visited on 24/08/2018).

[83] P. Ghimire, S. Bęczkowski, S. Munk-Nielsen, B. Rannestad and P. B. Thøgersen. "A review on real time physical measurement techniques and their attempt to predict wear-out status of IGBT". In: *Proceedings of the 15th European Conference on Power Electronics and Applications (EPE)* (Lille, France, 2–6 Sept. 2013). 2013, pp. 1–10. ISBN: 978-1-4799-0116-6. DOI: 10.1109/EPE.2013.6634419.

[84] S. Dusmez, S. H. Ali, M. Heydarzadeh, A. S. Kamath, H. Duran and B. Akin. "Aging Precursor Identification and Lifetime Estimation for Thermally Aged Discrete Package Silicon Power Switches". In: *IEEE Transactions on Industry Applications* 53.1 (Jan. 2017), pp. 251–260. ISSN: 0093-9994. DOI: 10.1109/TIA.2016.2603144.

[85] S. Dusmez and B. Akin. "An accelerated thermal aging platform to monitor fault precursor on-state resistance". In: *Proceedings of the 2015 IEEE International Electric Machines & Drives Conference (IEMDC)* (Coeur d'Alene, ID, USA, 10–13 May 2015). 2015, pp. 1352–1358. ISBN: 978-1-4799-7941-7. DOI: 10.1109/IEMDC.2015.7409238.

[86] J. Celaya, A. Saxena, P. Wysocki, S. Saha and K. Goebel. "Towards prognostics of power mosfets: Accelerated aging and precursors of failure". In: *Proceedings of the Annual Conference of the Prognostics and Health Management Society 2010 (PHM)* (Portland, OR, USA, 13–16 Oct. 2010). 2010. ISBN: 978-1-936263-01-1. URL: https://www.phmsociety.org/sites/phmsociety.org/files/phm_submission/2010/phmc_10_009.pdf (visited on 24/08/2018).

[87] L. Dupont, S. Lefebvre, M. Bouaroudj, Z. Khatir, J. Faugieres and F. Emorine. "Ageing Test Results of low voltage MOSFET Modules for electrical vehicles". In: *Proceedings of the 2007 European Conference on Power Electronics and Applications (EPE)* (Aalborg, Denmark, 2–5 Sept. 2007). 2007, pp. 1–10. ISBN: 978-92-75815-10-8. DOI: 10.1109/EPE.2007.4417433.

[88] H. Oh, B. Han, P. McCluskey, C. Han and B. D. Youn. "Physics-of-Failure, Condition Monitoring, and Prognostics of Insulated Gate Bipolar Transistor Modules: A Review". In: *IEEE Transactions on Power Electronics* 30.5 (May 2015), pp. 2413–2426. ISSN: 0885-8993. DOI: 10.1109/TPEL.2014.2346485.

[89] D. Xiang, L. Ran, P. Tavner, A. Bryant, S. Yang and P. Mawby. "Monitoring Solder Fatigue in a Power Module Using Case-Above-Ambient Temperature Rise". In: *IEEE Transactions on Industry Ap-*

*plications* 47.6 (Nov.–Dec. 2011), pp. 2578–2591. ISSN: 0093-9994. DOI: 10.1109/TIA.2011.2168556.

[90] L. Feller, S. Hartmann and D. Schneider. "Lifetime analysis of solder joints in high power IGBT modules for increasing the reliability for operation at 150 °C". In: *Microelectronics Reliability* 48.8-9 (Aug.–Sept. 2008). 19th European Symposium on Reliability of Electron Devices, Failure Physics and Analysis (ESREF 2008), pp. 1161–1166. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2008.07.019.

[91] U. Karki and F. Peng. "Effect of Gate Oxide Degradation on Electrical Parameters of Power MOSFETs". In: *IEEE Transactions on Power Electronics* (Feb. 2018). ISSN: 0885-8993. DOI: 10.1109/TPEL.2018.2801848.

[92] C. Chen, X. Ye, Y. Wang, J. Xu and G. Zhai. "PHM application of power converters using health precursor of power MOSFETs". In: *Proceedings of the 2015 Prognostics and System Health Management Conference (PHM)* (Beijing, China, 21–23 Oct. 2015). 2015, pp. 1–5. ISBN: 978-1-4673-8554-1. DOI: 10.1109/PHM.2015.7380078.

[93] Z. Pavlovic, I. Manic, Z. Prijic, V. Davidovic and N. Stojadinovic. "Influence of gate oxide charge density on VDMOS transistor ON-resistance". In: *Proceedings of the 22nd International Conference on Microelectronics (ICMEL)* (Nis, Yugoslavia, 14–17 July 2000). Vol. 2. 2000, pp. 663–666. ISBN: 0-7803-5235-1. DOI: 10.1109/ICMEL.2000.838777.

[94] J. R. Celaya, A. Saxena, C. S. Kulkarni, S. Saha and K. Goebel. "Prognostics approach for power MOSFET under thermal-stress aging". In: *Proceedings of the 2012 Annual Reliability and Maintainability Symposium (RAMS)* (Reno, NV, USA, 23–26 Jan. 2012). 2012, pp. 1–6. ISBN: 978-1-4577-1851-9. DOI: 10.1109/RAMS.2012.6175487.

[95] F. Erturk and B. Akin. "A method for online ageing detection in SiC MOSFETs". In: *2017 IEEE Applied Power Electronics Conference and Exposition (APEC)*. Mar. 2017, pp. 3576–3581. DOI: 10.1109/APEC.2017.7931211.

[96] N. Baker, H. Luo and F. Iannuzzo. "Simultaneous on-state voltage and bond-wire resistance monitoring of silicon carbide MOSFETs". In: *Energies* 10.3 (Mar. 2017). DOI: 10.3390/en10030384.

[97] J. Due, S. Munk-Nielsen and R. Nielsen. "Lifetime investigation of high power IGBT modules". In: *Proceedings of the 14th European Conference on Power Electronics and Applications (EPE)* (Birmingham, UK, 30 Aug.–1 Sept. 2011). 2011, pp. 1–8. ISBN: 978-90-75815-15-3. URL: `https://ieeexplore-ieee-org.libproxy.aalto.fi/stamp/stamp.jsp?tp=&arnumber=6020272` (visited on 25/08/2018).

[98] F. Stella, G. Pellegrino, E. Armando and D. Daprà. "On-line temperature estimation of SiC power MOSFET modules through on-state resistance mapping". In: *Proceedings of the 2017 IEEE Energy Conversion Congress and Exposition (ECCE)* (Cincinnati, OH, USA, 1–5 Oct. 2017). 2017, pp. 5907–5914. ISBN: 978-1-5090-2998-3. DOI: `10.1109/ECCE.2017.8096976`.

[99] N. Stojadinovic, I. Manic, S. Djoric-Veljkovic, V. Davidovic, S. Golubovic and S. Dimitrijev. "Mechanisms of positive gate bias stress induced instabilities in power VDMOSFETs". In: *Microelectronics Reliability* 41.9 (Sept.–Oct. 2001), pp. 1373–1378. ISSN: 0026-2714. DOI: `10.1016/S0026-2714(01)00143-3`.

[100] R. Green, A. Lelis and D. Habersat. "Threshold-voltage bias-temperature instability in commercially-available SiC MOSFETs". In: *Japanese Journal of Applied Physics* 55.4S (2016). URL: `http://stacks.iop.org/1347-4065/55/i=4S/a=04EA03` (visited on 27/08/2008).

[101] A. J. Lelis, R. Green, D. B. Habersat and M. El. "Basic Mechanisms of Threshold-Voltage Instability and Implications for Reliability Testing of SiC MOSFETs". In: *IEEE Transactions on Electron Devices* 62.2 (Feb. 2015), pp. 316–323. ISSN: 0018-9383. DOI: `10.1109/TED.2014.2356172`.

[102] T. T. Nguyen, A. Ahmed, T. V. Thang and J. H. Park. "Gate Oxide Reliability Issues of SiC MOSFETs Under Short-Circuit Operation". In: *IEEE Transactions on Power Electronics* 30.5 (May 2015), pp. 2445–2455. ISSN: 0885-8993. DOI: `10.1109/TPEL.2014.2353417`.

[103] T. Santini, S. Morand, M. Fouladirad, L. Phung, F. Miller, B. Foucher, A. Grall and B. Allard. "Accelerated degradation data of SiC MOSFETs for lifetime and Remaining Useful Life assessment". In: *Microelectronics Reliability* 54.9-10 (Sept.–Oct. 2014), pp. 1718–1723. ISSN: 0026-2714. DOI: `10.1016/j.microrel.2014.07.082`.

[104] X. Ye, C. Chen, Y. Wang, G. Zhai and G. J. Vachtsevanos. "Online Condition Monitoring of Power MOSFET Gate Oxide Degradation Based on Miller Platform Voltage". In: *IEEE Transactions on Power*

*Electronics* 32.6 (June 2017), pp. 4776–4784. ISSN: 0885-8993. DOI: 10.1109/TPEL.2016.2602323.

[105] U. Karki and F. Z. Peng. *Precursors of gate oxide degradation in SiC power MOSFETs*. Feb. 2018. arXiv: 1802.08085 [physics.gen-ph]. URL: https://arxiv.org/abs/1802.08085 (visited on 17/07/2018).

[106] A. Fayyaz and A. Castellazzi. "High temperature pulsed-gate robustness testing of SiC power MOSFETs". In: *Microelectronics Reliability* 55.9-10 (Aug.–Sept. 2015), pp. 1724–1728. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2015.06.141.

[107] R. Ouaida, M. Berthou, J. León, X. Perpiñà, S. Oge, P. Brosselard and C. Joubert. "Gate Oxide Degradation of SiC MOSFET in Switching Conditions". In: *IEEE Electron Device Letters* 35.12 (Dec. 2014), pp. 1284–1286. ISSN: 0741-3106. DOI: 10.1109/LED.2014.2361674.

[108] A. Ortiz-Conde, F. J. García-Sánchez, J. Muci, A. T. Barrios, J. J. Liou and C.-S. Ho. "Revisiting MOSFET threshold voltage extraction methods". In: *Microelectronics Reliability* 53.1 (Jan. 2013), pp. 90–104. ISSN: 0026-2714. DOI: 10.1016/j.microrel.2012.09.015.

[109] *Power Tester 1500A - Lifetime Testing and Failure Diagnosis of High-Power Semiconductors*. Mentor. 2014. URL: http://s3.mentor.com/public_documents/datasheet/products/mechanical/power-tester-1500a-ds.pdf (visited on 26/08/2018).

[110] V. Szekely, S. Torok, E. Nikodemusz, G. Farkas and M. Rencz. "Measurement and evaluation of thermal transients". In: *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference (IMTC)* (Budapest, Hungary, 21–23 May 2001). Vol. 1. 2001, pp. 210–215. ISBN: 0-7803-6646-8. DOI: 10.1109/IMTC.2001.928814.

[111] T. Funaki and S. Fukunaga. "Difficulties in characterizing transient thermal resistance of SiC MOSFETs". In: *Proceedings of the 22nd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)* (Budapest, Hungary, 21–23 Sept. 2016). 2016, pp. 141–146. ISBN: 978-1-5090-5451-0. DOI: 10.1109/THERMINIC.2016.7749042.

[112] W. Wu, M. Held, P. Jacob, P. Scacco and A. Birolini. "Investigation on the long term reliability of power IGBT modules". In: *Proceedings of the 1995 International Symposium on Power Semiconductor Devices and ICs (ISPSD)* (Yokohama, Japan, 23–25 May 1995). 1995, pp. 443–448. ISBN: 0-7803-2618-0. DOI: 10.1109/ISPSD.1995.515079.

[113] Z. Sarkany, W. He and M. Rencz. "Temperature change induced degradation of SiC MOSFET devices". In: *Proceedings of the 15th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)* (Las Vegas, NV, USA, 31 May–3 June 2016). 2016, pp. 1572–1579. ISBN: 978-1-4673-8121-5. DOI: 10.1109/ITHERM.2016.7517736.

[114] G. Sonnenfeld, K. Goebel and J. R. Celaya. "An agile accelerated aging, characterization and scenario simulation system for gate controlled power transistors". In: *Proceedings of the 2008 IEEE AUTOTESTCON* (Salt Lake City, UT, USA, 8–11 Sept. 2008). 2008, pp. 208–215. ISBN: 978-1-4244-2225-8. DOI: 10.1109/AUTEST.2008.4662613.

# Appendix A

## Power cycling settings and results

Table A.1: Power cycling parameter settings.

| ID | Type | Group | $t_{on}/s$ | $t_{off}/s$ | $I_{cycle}/A$ | $I_{sense}/A$ | $V_{G_+}/V$ | $V_{G_-}/V$ |
|---|---|---|---|---|---|---|---|---|
| 000 | A | 1 | 2 | 2 | 9.00 | -1 | 18 | -3 |
| 001 | A | 1 | 2 | 2 | 9.75 | -1 | 18 | -3 |
| 002 | B | 1 | 2 | 2 | 10.5 | -1 | 18 | -3 |
| 003 | A | 1 | 2 | 2 | 9.75 | -1 | 18 | -3 |
| 004 | A | 1 | 2 | 2 | 9.75 | -1 | 18 | -3 |
| 005 | A | 1 | 2 | 2 | 9.75 | -1 | 18 | -3 |
| 006 | B | 1 | 2 | 2 | 10.25 | -1 | 18 | -3 |
| 011 | A | 3 | 5 | 10 | 9.85 | -1 | 18 | -5 |
| 012 | A | 3 | 5 | 10 | 9.85 | -1 | 18 | -5 |
| 013 | B | 3 | 5 | 10 | 10.45 | -1 | 18 | -6 |
| 015 | B | 3 | 5 | 10 | 10.45 | -1 | 18 | -6 |

Table A.2: Power cycling stress characteristics and resulted life.

| ID | Type | Group | $T_{j_{max}}$/°C* | $\Delta T_j$/°C* | Cycles to failure/- |
|---|---|---|---|---|---|
| 000 | A | 1 | 133.15 | 83.73 | 8159 |
| 001 | A | 1 | 138.63 | 83.72 | 8554 |
| 002 | B | 1 | 148.03 | 96.19 | 54978 |
| 003 | A | 1 | 128.20 | 77.07 | 176626 |
| 004 | A | 1 | 141.29 | 88.69 | 54647 |
| 005 | A | 1 | 134.55 | 83.63 | 60142 |
| 006 | B | 1 | 132.18 | 92.50 | †674172 |
| 011 | A | 3 | 144.54 | 119.76 | 13458 |
| 012 | A | 3 | 143.16 | 118.69 | 31066 |
| 013 | B | 3 | 149.37 | 126.57 | 11819 |
| 015 | B | 3 | 161.78 | 136.83 | 30787 |

*The average value at the start of the test. Calculated by skipping the first 10 and averaging the subsequent 40 data points.
†Did not fail.