THE UNIVERSITY OF WARWICK

Library Declaration and Deposit Agreement

1. STUDENT DETAILS

Please complete the following:

Full name: ……Hui Huang……

University ID number: ……0756774……

2. THESIS DEPOSIT

2.1 I understand that under my registration at the University, I am required to deposit my thesis with the University in BOTH hard copy and in digital format. The digital version should normally be saved as a single pdf file.

2.2 The hard copy will be housed in the University Library. The digital version will be deposited in the University's Institutional Repository (WRAP). Unless otherwise indicated (see 2.3 below) this will be made openly accessible on the Internet and will be supplied to the British Library to be made available online via its Electronic Theses Online Service (EThOS) service.
[At present, theses submitted for a Master's degree by Research (MA, MSc, LLM, MS or MMedSci) are not being deposited in WRAP and not being made available via EthOS. This may change in future.]

2.3 In exceptional circumstances, the Chair of the Board of Graduate Studies may grant permission for an embargo to be placed on public access to the hard copy thesis for a limited period. It is also possible to apply separately for an embargo on the digital version. (Further information is available in the *Guide to Examinations for Higher Degrees by Research*.)

2.4 *If you are depositing a thesis for a Master's degree by Research, please complete section (a) below. For all other research degrees, please complete both sections (a) and (b) below:*

(a) Hard Copy

I hereby deposit a hard copy of my thesis in the University Library to be made publicly available to readers (please delete as appropriate) ~~EITHER immediately OR after an embargo period of …….. months/years as agreed by the Chair of the Board of Graduate Studies.~~

I agree that my thesis may be photocopied. YES / ~~NO~~ (*Please delete as appropriate*)

(b) Digital Copy

I hereby deposit a digital copy of my thesis to be held in WRAP and made available via EThOS.

Please choose one of the following options:

EITHER My thesis can be made publicly available online. YES / ~~NO~~ (*Please delete as appropriate*)

OR My thesis can be made publicly available only after…..[date] (Please give date)
~~YES~~ / NO (*Please delete as appropriate*)

OR My full thesis cannot be made publicly available online but I am submitting a separately identified additional, abridged version that can be made available online.
~~YES~~ / NO (*Please delete as appropriate*)

OR My thesis cannot be made publicly available online. ~~YES~~ / NO (*Please delete as appropriate*)

3. GRANTING OF NON-EXCLUSIVE RIGHTS

Whether I deposit my Work personally or through an assistant or other agent, I agree to the following:

Rights granted to the University of Warwick and the British Library and the user of the thesis through this agreement are non-exclusive. I retain all rights in the thesis in its present version or future versions. I agree that the institutional repository administrators and the British Library or their agents may, without changing content, digitise and migrate the thesis to any medium or format for the purpose of future preservation and accessibility.

4. DECLARATIONS

(a) I DECLARE THAT:

- I am the author and owner of the copyright in the thesis and/or I have the authority of the authors and owners of the copyright in the thesis to make this agreement. Reproduction of any part of this thesis for teaching or in academic or other forms of publication is subject to the normal limitations on the use of copyrighted materials and to the proper and full acknowledgement of its source.

- The digital version of the thesis I am supplying is the same version as the final, hard-bound copy submitted in completion of my degree, once any minor corrections have been completed.

- I have exercised reasonable care to ensure that the thesis is original, and does not to the best of my knowledge break any UK law or other Intellectual Property Right, or contain any confidential material.

- I understand that, through the medium of the Internet, files will be available to automated agents, and may be searched and copied by, for example, text mining and plagiarism detection software.

(b) IF I HAVE AGREED (in Section 2 above) TO MAKE MY THESIS PUBLICLY AVAILABLE DIGITALLY, I ALSO DECLARE THAT:

- I grant the University of Warwick and the British Library a licence to make available on the Internet the thesis in digitised format through the Institutional Repository and through the British Library via the EThOS service.

- If my thesis does include any substantial subsidiary material owned by third-party copyright holders, I have sought and obtained permission to include it in any version of my thesis available in digital format and that this permission encompasses the rights that I have granted to the University of Warwick and to the British Library.

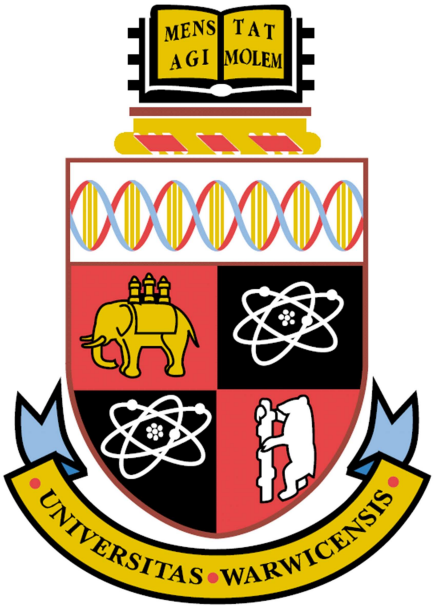
5. LEGAL INFRINGEMENTS

I understand that neither the University of Warwick nor the British Library have any obligation to take legal action on behalf of myself, or other rights holders, in the event of infringement of intellectual property rights, breach of contract or of any other right, in the thesis.

Please sign this agreement and return it to the Graduate School Office when you submit your thesis.

Student's signature: .....…… Date: .....

Lifetime Prediction for Power Converters



Hui Huang

School of Engineering

University of Warwick

Dissertation submitted for the degree of

Doctor of Philosophy

July 2012

Declaration

The work presented in this thesis was carried out in the School of Engineering, University of Warwick, during the period October 2007 to August 2011 under the supervision of Prof. Philip Mawby.

The author wishes to declare that apart from commonly understood and accepted ideas, or where reference is made to the work of others, the work in this thesis is his own. It has not been submitted in part, or in whole, to any other university for a degree, diploma or other qualification.

H. Huang
July, 2012

"*To my loving parents, who offered me unconditional love and provided me with the foundation to stand upon to reach for success*"

"*To Yao, who has been a great source of motivation and inspiration, for her love which I cannot live without, for her patience and support which I will forever grateful*"

Acknowledgements

Firstly, I would like to my supervisors Prof. Philip Mawby for his guidance and support through out my PhD studies. His advice, encouragement and feedback is invalueable for me and I shall be forever grateful.

I would also thank Engineering and Physical Sciences Research Council (EPSRC) for their financial support of this project.

I must thank my colleagues within the PEATER group, especially Benedict Donnellan for his numerous help with my writing, Dr. C.F. Tong for his useful advices and overall help, Dr. K. K. Leong for his help on proofreading of this work and daily chat, Dr. Shaoyong Yang for his cooperation on building the power cycling rig, Dr. M. R. Jennings for his cooperation on building the gate oxide reliability test rig, Dr. D. Hamilton for his useful discussion and advices on power module packaging and reliability, Dr. Nii-Adotei Parker-Allotey for his help in the Matlab simulation. I would also like to thank Stewart Edris and his fellow technicians for their helpful advices and excellent workmanship.

Finally, I would like to thank all my friends both within the University of Warwick and elsewhere, for their support and making them an invalueable experience.

Abstract

Renewable energy is developing rapidly and gaining more and more commercial viability. High reliability of the generation system is essential to maximize the output power. The power inverter is an important unit in this system and is believed to be one of the most unreliable parts. In the case of wind power generation, especially in off-shore wind, when the system reliability requirement is high, a technique to predict the inverter lifetime is invaluable as it would help the inverter designer optimize his design for minimal maintenance.

Previous researchers studying inverter lifetime prediction, focus either at device level such as device fatigue damage models, or at system level which require experimental data for their selected device. This work presents a new method to estimate the inverter lifetime from a given mission profile within a reasonable simulation time. Such model can be used as a converter design tool or an on-line lifetime estimation tool after being configured to a real converter system.

The key contribution of this work is to link the physics of the power devices to a large scale system simulation within a reasonable framework of time. With this technique, the system down time can be reduced and therefore more power can be generated. Also, the failure damage to the system is avoided which reduces the maintenance cost. A power cycling test is designed to gather the lifetime data of a selected IGBT module. Die-attach solder fatigue is found out to be the dominant failure mode of this IGBT module. The accuracy of widely accepted Miner's rule, which accumulates damage linearly, is discussed and a nonlinear accumulation method is promoted to predict the lifetime of power inverters.

Contents

List of Figures

List of Tables

Nomenclature

ADE	Ambipolar Diffusion Equation
AF	Accelerated factor
Al	Aluminum
CJC	Cold junction compensation
CNC	Current number of cycle
CSR	Carrier storage region
CTE	Coefficient of thermal expansion (ppm/K)
DCB	Direct copper bonding
DFIG	Doubly-fed induction generator
DMH	Dead man's handle
GaN	Gallium nitride
HVDC	High voltage direct current
IC	Infrared camera
IGBT	Insulated-gate bipolar transistor
K	Kelvin
MOSFET	Metal-oxide semiconductor field effect transistor
MTBF	Mean time between failure
MTTR	Mean time to repair
NPT	Non-punch-through
PDEs	Partial differential equations
PM	Permanent magnet
PT	Punch-through
PV	Photovoltaics
Si	Silicon
SiC	Silicon carbide
SPWM	Sin-triangle pulse width modulation
SSE	Sum of squared error
SVPWM	Space vector pulse width modulation
TDDB	Time dependent dielectric breakdown

TSP	Temperature sensitive parameter
USD	United State dollar

α	Weibull scale parameter
α_{Al}	CTE of Al (ppm/K)
α_{Si}	CTE of Si (ppm/K)
β	Weibull shape parameter
Δa_{el}	Elastic deformation (m)
Δa_{pl}	Plastic deformation (m)
Δn	Excess electron density (cm^{-3})
Δp	Excess hole density (cm^{-3})
ϵ	Permittivity of silicon ($1.04 \times 10^{12}\ F \cdot cm^{-1}$)
ϵ_{el}	Elastic strain
ϵ_{pl}	Plastic strain
λ	Short channel parameter (V^{-1})
λ_{th}	Heat conductance
$\mathscr{E}$	Electric filed (V/m)
$\mathscr{E}_d$	Electric field in drift region (V/m)
$\mathscr{E}_{BD}$	Breakdown electric filed (V/m)
μ_i	Mobility due to ionized impurities scattering ($m^2/(V \cdot s)$)
μ_l	Mobility due to acoustic phonon scattering ($m^2/(V \cdot s)$)
μ_n	Electron mobility ($m^2/(V \cdot s)$)
μ_p	Hole mobility ($m^2/(V \cdot s)$)
ψ_c	Contact potential (eV)
ρ	Density of the material
$\rho(x)$	Total charge density distribution along x coordinate (cm^{-3})
ρ_{water}	Density of water
σ	Stress (N/m)
σ_n	Electron capture cross section (cm^2)
σ_p	Hole capture cross section (cm^2)
σ_y	Yeild strength (MPa)
τ_{BF}	Carrier lifetime in N buffer layer (PT) (μs)
τ_{bHL}	High-level injection lifetime due to band-to-band recombination (s)
τ_{bn}	Electron lifetime due to band-to-band recombination (s)
τ_{bp}	Hole lifetime due to band-to-band recombination (s)
τ_{tHL}	High-level injection lifetime due to bulk trap recombination (s)
τ_{tn}	Electron lifetime due to bulk trap recombination (s)
τ_{tp}	Hole lifetime due to bulk trap recombination (s)
ε	Dielectric constant of Si (1.04×10^{-12} F/cm)
c	Thermal capacitance

C_{dep}	Depletion layer capacitance (F)
C_{GC}	IGBT gate-collector capacitance / Miller capacitance (F)
C_{GE}	IGBT gate-emitter capacitance (F)
C_{ies}	IGBT input capacitance (nF)
C_{oes}	IGBT output capacitance (nF)
C_{OX}	Gate oxide capacitance (F)
C_{res}	IGBT reverse capacitance (nF)
C_{th}	Thermal capacitance
C_{water}	Thermal capacitance of water
D	Damage
D	Duty ratio of cyclic load current
D_n	Electron diffusion constant (m^2/s)
D_p	Hole diffusion constant (m^2/s)
E	Youngs modulus (GPa)
E_a	Activation energy (eV)
E_c	Conduction bands J
E_c	Critical electrical field value for silicon (V/m)
E_g	Bandgap J
E_i	Fermi level energy (J)
E_t	Bulk trap energy in Si (J)
E_v	Valence bands J
E_{Fn}	Electron quasi Fermi level (J)
E_{Fp}	Hole quasi Fermi level (J)
E_{off}	IGBT turn-off loss (mJ)
E_{on}	IGBT turn-on loss (mJ)
f_s	Sampling frequency
f_e	Error figure between simulation results and experimental waveforms
G	Shear modulus (GPa)
G_{th}	Thermal generation rate
h_n	Electron recombination coefficient ($cm^4 \cdot s^{-1}$)
h_p	Hole recombination coefficient ($cm^4 \cdot s^{-1}$)
I_F	PN junction forward current (A)
I_g	IGBT gate current (A)
I_L	Load current (A)
$I_L[n]$	Instantaneous load current at sampling point [n] (A)
I_s	Reverse saturation current (A)
I_{cH}	Defined upper limit of load current during heating phase in power cycling test
I_{cL}	Defined lower limit of load current during heating phase in power cycling test
I_{CM}	Peak collector current from the RBSOA curve given in the datasheet (A)
$I_{ref}[n]$	Instantaneous reference current at sampling point [n] (A)
I_{RR}	Reverse recovery current (A)

I_{sne}	Minority carrier saturation current (PT) (A)
J_con	Conduction current density (A/m^2)
$J_diffusion$	Diffusion current density (A/m^2)
J_drift	Drift current density (A/m^2)
J_n	Electron current density (A/m^2)
J_p	Hole current density (A/m^2)
J_s	Reverse saturation current density (A/m^2)
k	Boltzmann's constant, 1.38×10^{-23} J/K
K_{PL}	MOS transconductance coefficient (A/V^2)
L_d	Diode series inductance (H)
L_e	IGBT Kelvin emitter inductance (mH)
L_n	Electron diffusion length (cm)
L_p	Hole diffusion length (cm)
L_S	Circuit stray inductance (H)
m_c	Conductivity effective mass (kg)
m_{water}	Mass of water
n	Electron density (cm^{-3})
n_0	Thermal equilibrium electron density (cm^{-3})
N_a	Ionized acceptor density (cm^{-3})
N_B	Drift region width (μm)
N_d	Ionized donor density (cm^{-3})
N_f	Number of cycle to failure
n_i	Intrinsic carrier density (cm^{-3})
N_P	Weibull percentile
N_t	Bulk trap density in Si (cm^{-3})
N_H	Doping concentration of N buffer layer (PT) (cm^{-3})
n_{p0}	Thermal equilibrium electron density in p-type Si (cm^{-3})
n_p	Electron density in p-type Si (cm^{-3})
p	Hole density (cm^{-3})
p_0	thermal equilibrium hole density (cm^{-3})
P_{jc}	Heat conducts from IGBT junction to case
p_{n0}	thermal equilibrium hole density in n-type Si (cm^{-3})
p_n	Hole density in n-type Si (cm^{-3})
P_{water}	Heat extracted by water
q	Magnitude of the electron charge, 1.6×10^{-19} C
Q_d	Stored charge in the drift region (C)
Q_F	Total excess carrier charge in the drift region (C)
Q_n	Total excess carrier charge in n-side (C)
Q_p	Total excess carrier charge in p-side Si (C)
Q_{RR}	Reverse recovery charge (C)
R	Resistance per unit length

r	Cycle ratio
R_b	Band-to-band recombination rate
R_g	Gate resistance (Ω)
r_y	Plastic zone width (m)
R_{bc}	Recombination coefficient due to band-to-band recombination
R_{jc}	Thermal resistance between IGBT junction and case
R_{SP}	Resistance of the P-well (Ω)
R_{th}	Thermal resistance
T	Absolute temperature (K)
T_c	IGBT case temperature
t_c	Time between collision (s)
T_j	IGBT junction temperature
T_m	Mean IGBT junction temperature in power cycling test
T_{cH}	Defined upper limit of IGBT case temperature in power cycling test
T_{cL}	Defined lower limit of IGBT case temperature in power cycling test
t_{RR}	Reverse recovery time (s)
T_{varG}	Average water temperature (℃)
T_{win}	Water temperature before flow into the heatsink
T_{wout}	Water temperature after flow out of the heatsink
U_b	Net band-to-band transition rate
U_t	Net bulk trap transition rate
V_d	Carrier drift velocity (m/s)
V_d	Voltage across the drift region (V)
V_F	PN junction forward voltage drop (V)
V_L	Load voltage (V)
V_O	Inverter output voltage (V)
V_{AK}	Diode anode-cathode voltage (V)
V_{BD}	Breakdown voltage (V)
V_{BR}	Breakdown voltage from datasheet plus 150 to 200 V typical margin for general IGBTs (V)
V_{ce}	IGBT on-state voltage drop
V_{CM}	Peak IGBT collector voltage during turn-off (V)
V_{DC}	DC supply voltage (V)
V_{GE}	IGBT gate-emitter voltage (V)
$V_{GG(off)}$	IGBT off-state gate voltage (V)
$V_{GG(on)}$	IGBT on-state gate voltage (V)
$V_{L\_ave}[n]$	Average grid voltage over the switching period [n, n+1] (V)
$V_{O\_ave}[n]$	Average inverter output voltage over the switching period [n, n+1] (V)
V_{ref}	Reference voltage vector
V_{TH}	IGBT threshold voltage (V)
v_{th}	Thermal velocity (m/s)

V_{water}	Speed of water flow
W_D	Depletion layer width (cm)
W_d	Drift region width (cm)
W_{Dn}	Depletion layer width in n-type region (cm)
W_{Dp}	Depletion layer width in p-type region (cm)
W_H	Width of N buffer layer (PT) (μm)
x_{meas}	Experimentally measured waveforms
x_{sim}	Simulated results
Z_{th}	Thermal impedance

Chapter

1 Introduction

1.1 Energy Security and Low-carbon Society

The development of human society is highly reliant on energy. Fig. 1.1 shows that in 2008, 78% of the worlds total energy consumption was generated from fossil fuels [5]. Throughout recent history, the demand and consumption of fossil fuels has lead to many problems, such as air pollution and global warming. Furthermore, fossil fuels are finite resources which eventually will dwindle and become too expensive. The key to overcome these problems is to find alternative energy sources which are inexhaustible and environmentally friendly.

Nuclear as an alternative option is becoming more and more important. 439 nuclear power plants were installed in 31 countries upto and including 2008 and generate 2.8% of the worlds energy [6]. Nuclear power plants use the controlled nuclear fission rods to heat water and produce steam to drive the turbine. However, the debate about whether nuclear should be used as an energy source grows with installed capacity. Proponents suggest that nuclear power is a sustainable energy source which produces no conventional air pollution, such as green-house gases. They claim that the risk of the nuclear leakage is small and can be reduced by improving the technology. Opponents believe that nuclear power plants

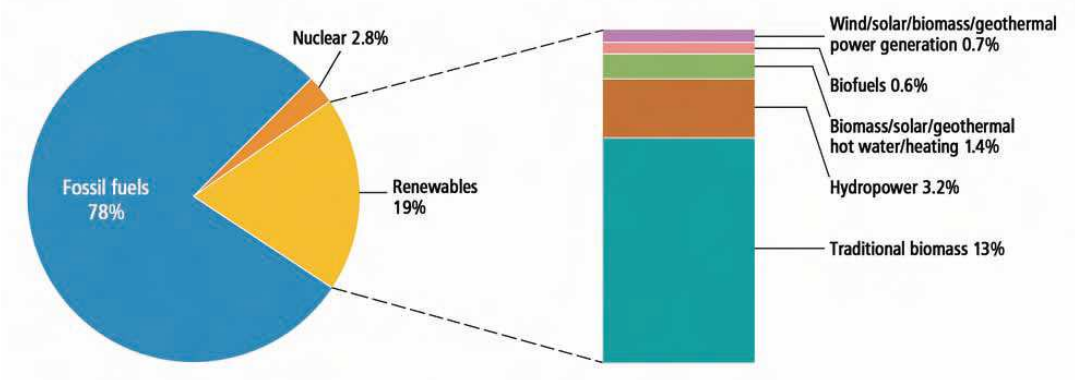


Figure 1.1: Energy consumption from different sources in 2008.

bring many threats to the people who live around them. In fact, there have been a number of serious nuclear accidents, for example the Chernobyl disaster in 1986, which caused a great deal of damage to the global environment and human health. The most recent nuclear accident is the Fukushima disaster in Japan on the 11^{th} March 2011, which cause huge damage to Japan's environment and economy. These accidents support the opponents' view and have caused a global recession in nuclear power.

Renewable energy is the energy which comes from natural sources, such as sunlight, wind, geothermal and tide which are naturally replenished. Therefore renewable energy does not have a generation limit and so are environmentally friendly. Furthermore, these energy sources could supply enough power for human society if they can be harnessed efficiently. Globally, about 1700TW of solar power is theoretically available over land for Photovoltaics (PV), which means even the capture of 1% of this power would supply more power than we currently need [6]. The available wind power globally is about 72TW, which is 5 times the worlds total power production [7]. However, the installed solar PV capacity was only 21GW and wind power capacity was 159GW in 2009 [5]. Fig. 1.2 shows some examples of renewable energy power plant. Clearly, renewable power plant can be built in desert or ocean where the environment is too severe for mankind to live. Therefore they will not affect the human's living space. In this case, high system reliability is required since the maintenance

(a) Solar energy power plant

(b) Offshore wind farm

Figure 1.2: Examples of renewable energy power plant

cost will be relevantly high, due to limited access.

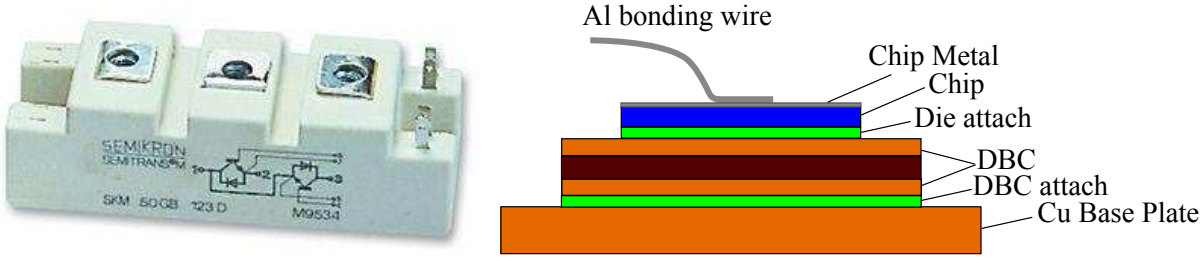
Although the market share of solar and wind energy only accounts for 0.7% of the total power consumption, it is growing very rapidly and gaining more and more attention from both researchers and business. Grid connected solar PV has grown by an average of 60% each year for the past decade while wind power capacity has grown at 27% annually for the past 5 years. The annual global investment in new renewable capacity reached 150 billion United State dollars (USD) in 2009 [6]. After the Kyoto summit in 1997, most of the worlds major countries committed to change the ways they used and supplied energy to limit the emission of green-house gases. Acts were passed by governments to legally constrict greenhouse gas emission, for example, UK's climate change act in 2008. However, the development of renewable energy is still highly government driven because the cost of energy generated from renewable sources is still higher than that from fossil fuels. Clearly, increasing the reliability of energy generation systems is a direct method of reducing the operation and maintenance cost. After all, the development potential of renewable energy is huge and more technical efforts is required.

1.2 The Role of Power Electronics in Energy Efficiency

The electrical energy from renewable sources normally cannot be despatched by the grid directly. In the case of wind power system, the frequency and voltage of the electricity generated by wind turbine changes as the wind speed varies. Energy conversion is required to modify the generated electricity to the grid frequency and voltage. Power converters are installed to accomplish this task because of their high efficiency, normally higher than 90%. Unlike classical electronics, power electronics process power rather than information. Nowadays, power electronic devices can be found in almost all electrical applications like mobile phone, traction, satellite, wind turbines and so on. At least 25% of all generated power passes through some form of power electronic system before being utilised [8].

Power electronic technology was originally developed at the beginning of 20^{th} century when the mercy arc valve, the first high power electronic device, was invented. Mercury arc valves were widely used to convert AC to DC until the 1960s when high voltage semiconductor devices like silicon diode and thyristor were developed. Metal Oxide Semiconductor Field-Effect Transistor (MOSFET) has become commercially available in 1976. With the advantage of both high switching frequency and controllable turn-off, they are widely used in many electrical applications. However, MOSFET is a unipolar device which faces the tradeoff between voltage blocking capability and on-state resistance. The insulated Gate Bipolar Transistor (IGBT) was invented to overcome this limitation. The voltage rating of an IGBT module, which is commercially available now, is up to 6500V. Fig. 1.2 shows the example of current available IGBT module and its basic structure.

After 100 years of development, the family of power electronic device contains a range of products, as shown in Fig 1.4. Thyristors are used in high power and low switching frequency applications such as the high voltage direct current (HVDC) Grid. It can be seen that the switching frequency of an IGBT is much higher than that of the thyristor therefore it is widely used in motor control applications. MOSFETs, whose switching frequency is even



(a) IGBT module (b) structure of packaging

Figure 1.3: IGBT module and different layers of packaging

higher, are typically used in the power supply for portable devices. In the case of renewable energy, higher switching frequency and power capacity devices are preferred to limit the harmonics and convert higher power levels. Due to these requirements, IGBT is preferred device for renewable applications.

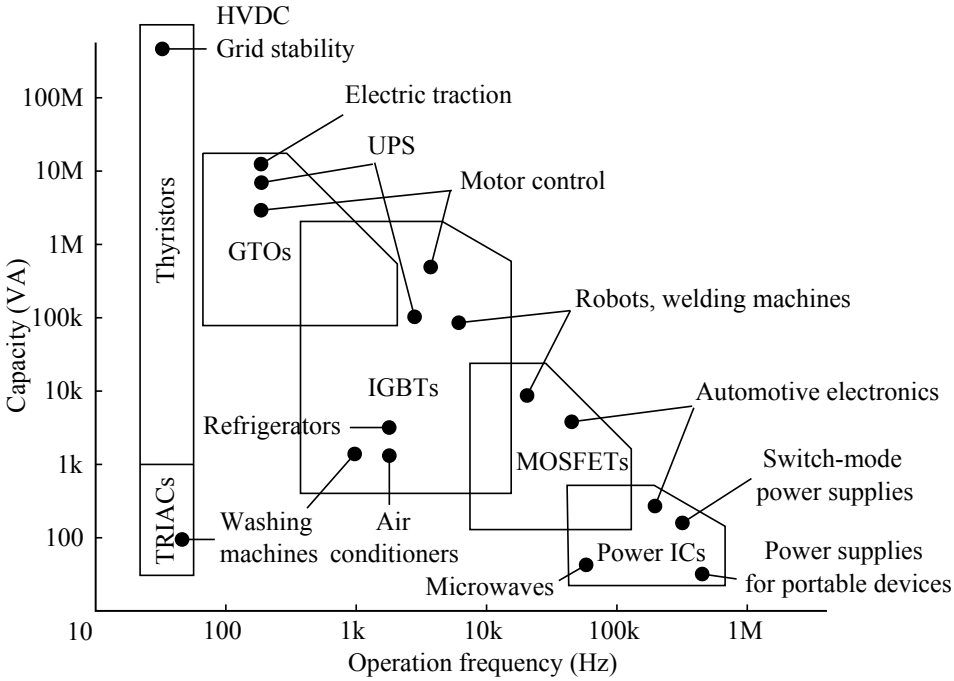


Figure 1.4: Typical example of power electronics and their applications.

1.3 Overview of Reliability Problems

Reliability is often defined as the ability of a device to fulfill its intended function during an interval of time [9]. Today, reliability engineering is applied to many products as failure can causes catastrophic consequences such as aircraft accidents. Even some small failures, such as the failure of a mobile phone, are a nuisance and could give the company that produces it a bad reputation. According to an industry based survey in 2009 [10], the converter is one of the most unreliable parts of an electrical system operating in a harsh environment, such as wind turbine. The cost of converter failure and maintenance is high in the case of off-shore operation due to the inaccessibility of the system. The system reliability can be greatly increased by replacing devices before they fail. Therefore the reliability of the converter is a critical issue.

The reliability of most products has often been represented by an idealised plot called a

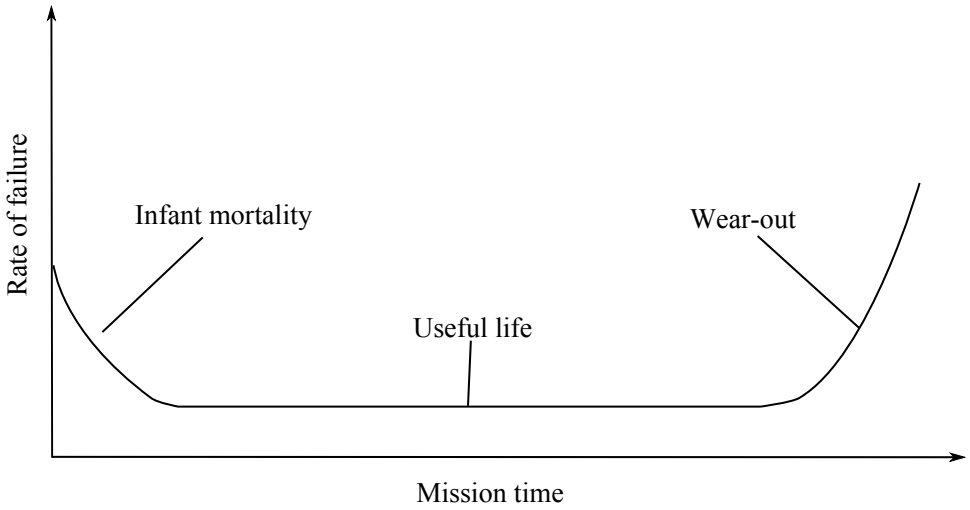


Figure 1.5: Idealized bathtub curve.

bathtub curve shows in Fig. 1.5. This curve consists of three regions: infant mortality, useful life and wear out. In the first region, failure is mainly caused by manufacturing defects or poor workmanship. The defective products tend to fail early, so the failure rate decreases with time. The failure rate reaches a constant level in the useful life region, where the failures observed are considered as random failures. In the last region, the products start wearing out and consequently the failure rate begins to rise as they experience more stresses.

Availability of a system over its life cycle is considered as a factor of system reliability. It is defined as the percentage of time when system is operational over its designed life cycle. Availability is the goal of most system users and could be increased by either improving reliability or reducing the Mean Time To Repair (MTTR) according to Eqn. 1.1. The early failure can be detected by the burn in test before leaving the factory to improve the device reliability. Estimating the devices lifetime, detecting their wear out region and then replacing the potential faulty devices could help optimise the maintenance to reduce the MTTR hence greatly increase the system availability, especially in the case of off-shore wind power generation as shown in Fig. 1.2 due to the inaccessibility of the system.

$$Availability = \frac{Mean\ Time\ Between\ Failure(MTBF)}{MTBF + MTTR} \tag{1.1}$$

Generally, there are four methods which are used to evaluate the lifetime of a power module [11]. They are the qualification procedure, theoretical calculations, field failure experiences, and the physics of failure method. Qualification procedure refers to a standard test with defined conditions. The test condition normally defined by international standard organisations, such as IEC, for a specific industry sector. The product that passes these tests are considered reliable. This method is simple for companies since no study is needed. However, this method normally only gives the general conditions and do not takes the application into account. Therefore its accuracy is not guaranteed.

Theoretical calculation refers to the estimation of reliability based on the accelerated tests and statistical models. The test results are fitted to a predefined statistical models and then extrapolated to the field conditions to give an estimation of the failure rate. The ratio of lifetime between the normal use level and higher test stress level is termed acceleration factor (AF) [12] . For the package failures, the lifetime is determined by cyclic thermomechanical stress, the Coffin-Manson equation is widely used to estimate the number of cycle to failure N_f, as shown in Eqn. 1.2. Where b and C are constants that fitted from accelerated test results. In this case, the AF could be given by Eqn. 1.3, where T_{jn} is the field IGBT junction temperature cycle and T_{ja} is the acceleration IGBT junction temperature cycle. This method is widely used to give the reliability estimation. However, its drawback is the difficulty in collecting data as well as the failure mechanisms at different operation condition might be different.

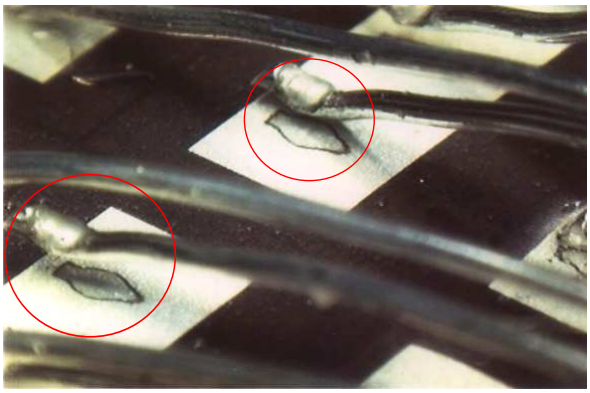
$$N_f = \left(\frac{C}{\Delta T_j}\right)^b \tag{1.2}$$

$$AF = \left(\frac{\Delta T_{ja}}{\Delta T_{jn}}\right)^b \tag{1.3}$$

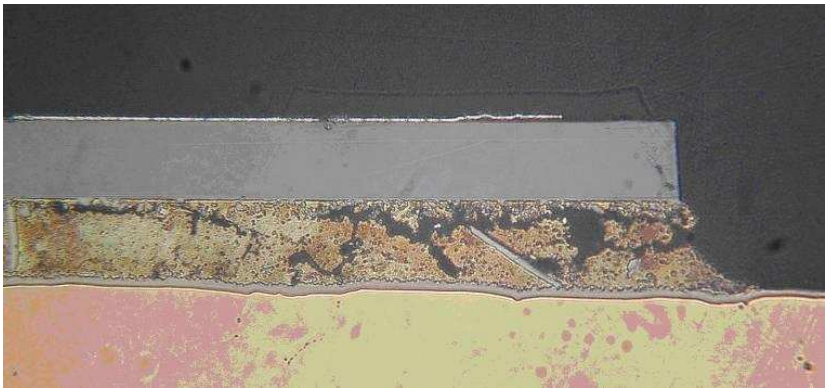
The field failure experience involves the failure data collection and analysis. Although this method gives accurate reliability estimation, it is unrealistic to carry out since the difficulty of data collection. Furthermore, it is not suitable for estimating the reliability of new technology since this means test the new product in field condition. For example a product with designed lifetime of 40 years, it is impossible to take this test.

The reliability estimation of the above three methods are based on the test or field data, no actual physics of failure mechanism is considered. While the physics of failure method [13, 14], which is developing rapidly in recent years, refers to the estimation of the device reliability based on the root cause of the failure process. This method requires a good understanding of the failure mechanisms, including the device material properties, the damage initiation, accumulation and propagation. Physics based failure models for different failure mechanisms are built to estimate the remaining lifetime of the device. In order to obtain an accurate reliability estimation in the field conditions, the mission profile is generated and used as the input of the lifetime damage simulation. The mission profile refers to a set of typical operational conditions that the device experienced during field operation.

There are various failure modes for IGBT modules, two of the most commonly observed ones are both packaging failures: bond wire damage, as shown in Fig. 1.6(a), and die-attach solder fatigue, as shown in Fig. 1.6(b). Fatigue is a process leading to the failure which is caused by repeated stress cycles below the tensile strength of the material [3]. Both failure modes are caused by the temperature cycling during operation and the thermal expansion coefficient (CTE) mismatch between adjacent layers. Therefore IGBT junction temperature T_j is one of the critical factors for lifetime estimation. However, T_j is difficult to measure during operation since the chip is buried in the module package. A possible option is to build an accurate electrothermal model to calculate the junction temperature during operation.



(a) Bond wire lift off [15]



(b) Die attach solder crack [16]

Figure 1.6: Two common failure mechanisms.

1.4 Scope of the Thesis

Previous researchers studying inverter lifetime prediction focus either at device level such as device fatigue damage models [4, 17], or at system level [18, 19], which lack of experimental reliability data for their selected device. This work investigates a new lifetime prediction framework, as shown in Fig. 1.7, to obtain an estimation of inverter lifetime in a large scale system simulation within a reasonable time estimation. Basically, there are three layers in this framework: device switching model, converter electrothermal model and lifetime damage model.

An introduction of some background knowledge is necessary in order to understand the

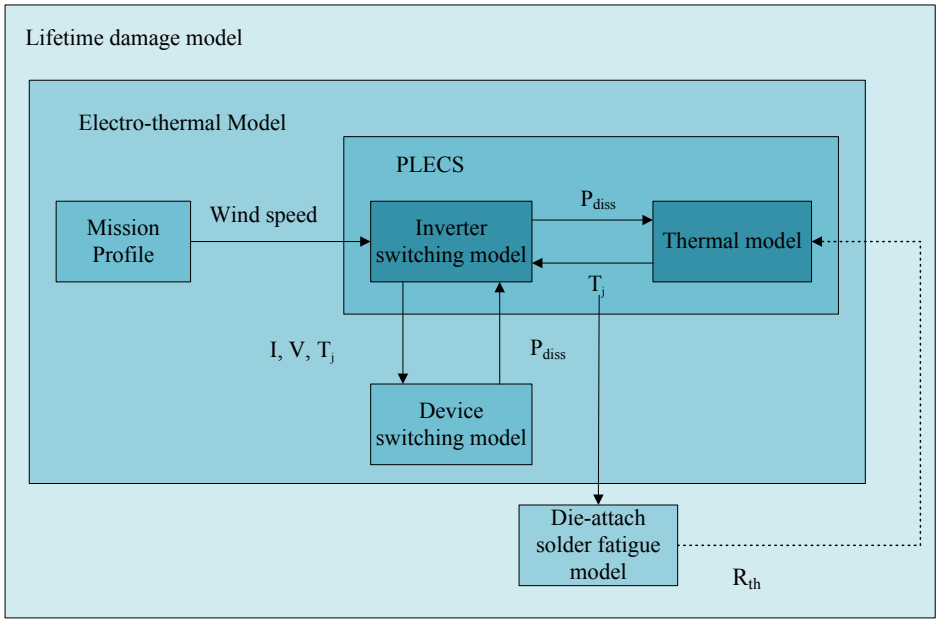


Figure 1.7: Framework of lifetime damage model.

lifetime damage model better. The device switching model used in this work is physics based Fourier series solution model [20, 21]. Therefore, chapter 2 introduces some basic physics of power semiconductor devices and their modelling techniques. Several basic problems are discussed, such as the principles of semiconductor devices, how temperature affect the behavior of the power device, and the principles of the physics based model.

Chapter 3 first introduces the current packaging technology of power modules. Different structures for the modern power module are introduced as well as their advantages compared with the older technology. Then it focuses on discussing the failure mechanisms of power modules. Die-attach solder fatigue and bond wire lift off is considered as the two main failure mechanisms that lead to the failure of a power module. The thermomechanical stress caused by the junction temperature cycling and the coefficient of thermal expansion (CTE) mismatch between adjacent layers is the main reason for the fatigue failure. The available models to express the reliability of these failure mechanisms are discussed in detail. Other failure mechanisms are briefly introduced at the end of this chapter.

Junction temperature is difficult to measure directly during inverter operation since the IGBT chip is buried in the power module. Therefore, it is desired to build a model that could simulate the junction temperature according to the operating conditions. Chapter 4 explains the detail of the inverter electrothermal model. In order to fit the physics based model to the selected IGBT module, a careful parameter extraction sequence is needed to ensure the simulation results match the experimental curves at different temperatures. A two-step parameter extraction method for the device models is introduced whose accuracy at different temperatures is shown in a previous paper [22]. Although the selected device switching model runs faster than finite element method (FEM) models, it still takes two to three seconds to simulate one switching event, which makes it impossible to apply directly in a large scale system simulation. Look-up-tables (LUTs) of the device switching loss, which can be generated by running the device switching model at different currents, voltages and junction temperatures, are used to overcome this problem. The IGBT and diode conduction loss LUTs are created from the device forward IV characteristics. During inverter simulation, these LUTs are interpolated to obtain the device power loss according to its operating conditions. Thermal models of the IGBT module and heatsink are extracted experimentally from their warming curves. IGBT junction temperature can then be obtained with the device power loss and thermal networks.

The lifetime data of most IGBT modules are not available publicly. Therefore, a power cycling rig is designed and built to test the selected modules and gather their lifetime information. Due to the differences between individual modules, multiple samples are required for each test condition. The IGBT on-state voltage drop (V_{ce}) and thermal resistance between IGBT junction and case (R_{jc}) are the indicators of module degradation, hence they are monitored throughout the power cycling test. Chapter 5 explains the principles of the power cycling test, which includes the design of the rig and analysis of the experimental data.

The module fatigue damage model fitted from power cycling data could only estimate

the device lifetime with a known constant junction temperature swing. However during field operation, it is impossible to keep the junction temperature cycles constant. To link the field junction temperature profile with the IGBT lifetime, a cycle counting method is needed to decompose the temperature profile into elementary cycles characterised by their amplitudes. Furthermore, a method to accumulate the damage for each temperature cycle is also required. Chapter 6 first introduces the weibull reliability analysis method and compares the fatigue damage model extracted from experimental data with other researcher's results. Then it reviews some cycle counting methods and discusses their pros and cons. Rainflow cycle counting method is selected and explained in detail. Different damage accumulation methods are introduced and discussed. Although Miner's law is widely used for design engineers, its accuracy is questionable in many applications. As fatigue damage accumulates, the thermal resistance of the solder joint increases, which will lead to a higher IGBT junction temperature and accelerate the damage process. Hence a feedback loop can be added to update the temperature profile regularly and improve the performance of the linear method, as shown by the doted line in Fig. 1.7. Both the power cycling results and the crack propagation theory indicates that the damage accumulation process accelerates when the device is heading towards the end of its lifetime. A nonlinear damage accumulation method first proposed by Marco and Starkey [23] is introduced. Its principle and validity are discussed in detail.

Chapter 7 describes an application of the inverter lifetime damage model in wind power system. The lifetime of the grid connected voltage source inverter is estimated according to a wind speed profile based on real time measurements. First, the topologies of wind power generation system are reviewed. The full rated power semiconductor scheme is selected for simplification. The control strategy of the inverter is then studied and discussed in detail. The simulated junction temperature profile is counted by the rainflow code and then accumulated by different accumulation methods. The Miner's law gives the least conservative results while the nonlinear method provides the most conservative estimation.

Finally, conclusions and possible future development are addressed in chapter 8.

Chapter

2 Power Electronics Physics and Modelling Technique

As discussed in chapter 1, T_j is a critical parameter for the lifetime estimation of power converters. However, T_j is difficult to measure directly during converter operation, since the IGBT chip is buried inside the module package. A possible approach is to build an accurate electrothermal model to calculate T_j from simulation. In order to obtain accurate simulation results under different operating conditions, physics based device models are selected to generate the switching power losses. Therefore it is essential to understand the principles of semiconductor devices. This chapter starts with the basic physics of power electronic devices, and then reviews the current available device models. Finally, the effects of temperature on the semiconductor devices are discussed.

2.1 Basic Physics of Power Electronic Devices

Silicon (Si) has been widely used in commercial electronic devices. Although new materials have been proposed to have better performance, such as silicon carbide (SiC) and gallium nitride (GaN), silicon devices still dominate the market because of their advantages in price and well developed production technology. Therefore, the power electronic devices mentioned

in this thesis refer to Si based devices and the device model used in this work is designed for Si based devices. The major properties of Si based devices are detailed in this section.

2.1.1 Conduction Processes in Silicon

Electrical current can flow in a material if there are charge carriers that are free to move when electric field is applied. Therefore the electrical conductivity of a material is determined by the density and mobility of its free carriers. This subsection will explain the free carrier creation and recombination mechanisms as well as their movements to form electrical current in Si.

2.1.1.1 Free carriers: Holes and Electrons

There is a forbidden energy range in which allowed states cannot exist for all semiconductors. The energy regions above the gap are called conduction bands (E_c) while the those below the gap are called valence bands (E_v). The energy difference between the lowest E_c and the highest E_v is called the bandgap (E_g), which is one of the most important parameters in semiconductor physics. Each Si atom has four outer electrons, which are normally shared with adjacent atoms to form covalent bonds. The bandgap can be thought of as energy needed to break the covalent bond. The covalent bonds could be broken by the energy carried by the Si atom due to its random thermal motion when temperature is above 0 K, which is known as thermal ionisation. The thermal energy of a atom increases as temperature rises hence the bandgap decreases in the same time. The variation of bandgap with temperature can be expressed by Eqn. 2.1 [24]. Where $E_g(0)$ is the bandgap at 0 K, a and b are constants given as: $E_g(0) = 1.169$ eV, $a = 4.9 \times 10^{-4}$ eV/K, $b = 655$ K. The bandgap for Si is 1.12 eV at room temperature.

$$E_g(T) = E_g(0) - \frac{aT^2}{T + b} \tag{2.1}$$

When a covalent bond breaks, a free hole as well as a free electron is created. In intrinsic Si, the thermal ionisation mechanism generates an equal number of electrons and holes. The thermal equilibrium density of electrons and holes, (n_i), is given by Eqn. 2.2 [25]. Where q is the magnitude of the electron charge, k is Boltzmann's constant, T is the absolute temperature and C is a constant of proportionality. At room temperature, $n_i \approx 10^{10} cm^{-3}$ in Si, which is too low for conducting electrical power efficiently. One advantage of semiconductor is that it can be doped with different types of impurities in various concentration to control its resistivity. The appropriate impurities for Si are elements from the third and fifth column of the periodic table, such as boron and phosphorous.

$$n_i^2 = C \cdot T^3 \cdot \exp(\frac{-qE_g}{kT}) \tag{2.2}$$

Fig. 2.1 shows the basic bond structures of Si doped with boron and phosphorous. Phosphorous has five valance electrons which are available for bonding with other atoms in a crystal. When it replaces one Si atom in a Si crystal, a negative charged electron is donated to the lattice and became a free carrier. Hence the phosphorous atom is called a donor. In the Si lattice, doped with phosphorous, electrons become more numerous than holes and are called majority carriers. This type of doped Si is termed n-type Si. Boron has only three valance electrons, thus an additional electron is needed to bond with four neighboring Si atoms when boron is introduced into a Si crystal. When a boron atom captures the needed electron from the Si lattice, an hole is created consequently. In this case, the electron concentration becomes less than the hole concentration and are called minority carriers. This type of Si is called p-type Si.

The doping concentration of impurities in semiconductor devices are normally between 10^{15} - 10^{18} cm^{-3}. Lightly doped n-type Si, which normally termed as n-, has a doping concentration between 10^{13} - 10^{14} cm^{-3}. This type of Si is commonly used as the drift region of high power devices. As the impurity density is much less than the density of Si atoms, its presence will not affect the thermal ionisation process in Si, which means the product of

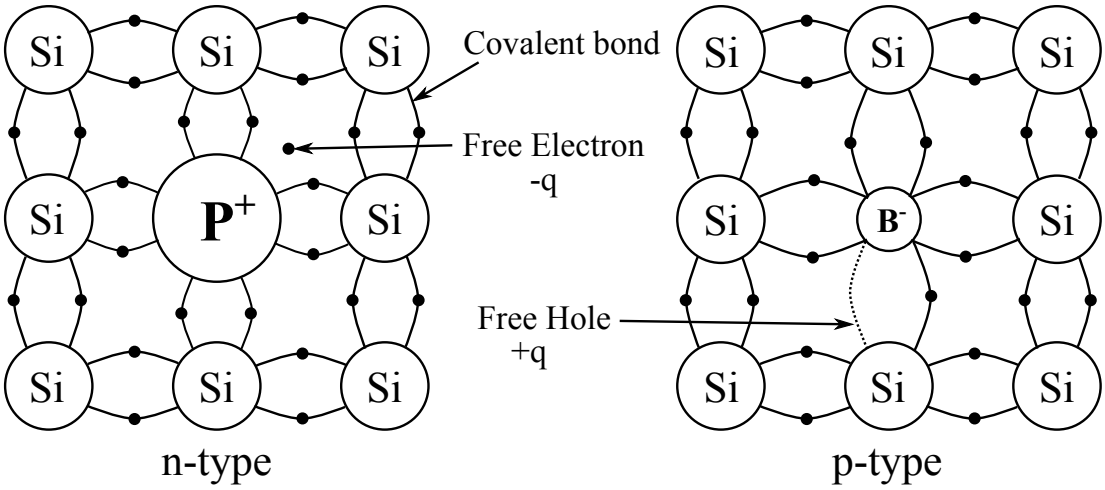


Figure 2.1: Silicon lattice dope with phosphorous (n-type) and boron (p-type).

the thermal equilibrium electron density (n_0) and the thermal equilibrium hole density (p_0) are still equal n_i^2, as shown in Eqn. 2.3.

$$p_0 n_0 = n_i^2 = C \cdot T^3 \cdot \exp(\frac{-qE_g}{kT}) \tag{2.3}$$

The positive charge density is the sum of p_0 and the ionised donor density N_d, while the negative charge density is the sum of n_0 and the ionised acceptor density N_a. Hence Eqn. 2.4 holds because of the space charge neutrality.

$$p_0 + N_d = n_0 + N_a \tag{2.4}$$

According to Eqns. 2.3 and 2.4, the electron and hole density can be calculated from the doping density. The intrinsic carrier density in Si lattice is lower than $10^{12}\ cm^{-3}$ when the temperature is less than 150 ℃ [26]. This is much lower than the impurity density. Hence the free carrier density can be considered approximate to the doping concentration. In the case of n-type Si, n_0 and p_0 can be estimated from Eqns. 2.5 and 2.6 respectively. Equivalent equations can be applied to p-type Si as well. From Eqn. 2.6, it is clear that the minority carrier density is proportional to the square of the intrinsic carrier density, hence is strongly

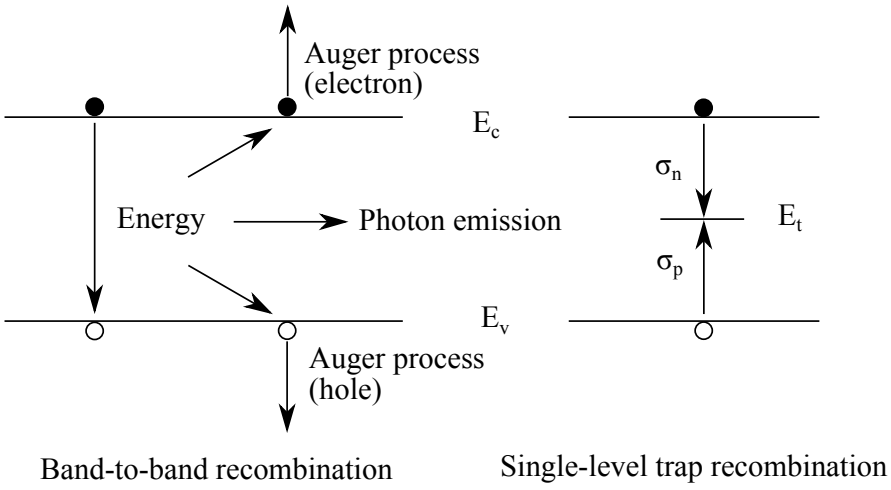


Figure 2.2: Recombination mechanisms.

temperature dependent.

$$n_0 \approx N_d \tag{2.5}$$

$$p_0 \approx \frac{n_i^2}{N_d} \tag{2.6}$$

2.1.1.2 Recombination

The last section introduced the free carrier creation mechanisms in Si, however, mechanisms for the disappearance of them must exist to maintain the thermal equilibrium condition. Such mechanisms are called recombination. Generally, there are two different mechanisms for free carriers to recombine, termed as band-to-band recombination and bulk trap recombination as shown in Fig. 2.2. Where σ_n and σ_p are the electron and hole capture cross sections respectively.

When an electron is displaced from the conduction band to the valence band and recombines with a hole, its energy can be conserved by emission of a photon or by transfer to another free carrier. The former process is called radiative process, while the latter is

termed Auger process. These two processes belong to band-to-band recombination. When a free carrier is captured by a bulk trap in the Si lattice, it also can be considered as a disappearance of free carrier. This process is termed bulk trap recombination. The density and energy of bulk traps are N_t and E_t respectively. E_t normally present within the bandgap, N_t is normally around 10^{14} cm^{-3} in Si lattice [26].

The band-to-band recombination rate R_b is proportional to the product of electron and hole densities as shown in Eqn. 2.7. where R_{bc} is the recombination coefficient which is highly temperature dependent. The typical value of R_{bc} for Si at room temperature is around 10^{-15} cm^3/s. The net band-to-band transition rate U_b is defined as the different between the band-to-band recombination and generation rate G_{th}, as shown in Eqn. 2.8. It is clear that U_b is zero at thermal equilibrium.

$$R_b = R_{bc} pn \tag{2.7}$$

$$U_b = R_b - G_{th} = R_{bc}(pn - n_i^2) \tag{2.8}$$

Under low-level injection, the excess carrier densities $\Delta p = \Delta n$ are lower than the majority carrier density which can be approximated as N_d in the case of a n-type Si. The minority density is then $p_n = p_{n0} + \Delta p$. Therefore, U_b can be approximated according to Eqn. 2.9, with the minority carrier lifetime defined by Eqn. 2.10. Equivalent equations holds in the case of p-type Si.

$$U_b \approx R_{bc} \Delta p N_d = \frac{\Delta p}{\tau_{bp}} \tag{2.9}$$

$$\tau_{bp} = \frac{1}{R_{bc} N_d} \tag{2.10}$$

Under high-level injection when the excess carrier densities are higher than the majority

carrier density, both the hole and electron densities can be approximated as the excess carrier densities. Hence the excess carrier lifetime can be estimated according to Eqn. 2.11. It is clear that the lifetime limited by the band-to-band recombination decreases as the excess carrier density rises. This can be explained as that the probability of the carrier interaction increases as its density rises.

$$\tau_{bHL} = \tau_{bn} = \tau_{bp} = \frac{1}{R_{bc}\Delta n} \tag{2.11}$$

The net bulk trap transition rate can be described by the Shockley-Read-Hall statistics as shown in Eqn. 2.12 [27]. Where p and n are the densities of hole and electron respectively, v_{th} is the thermal velocity which is given by Eqn. 2.13 [24], m_c is the conductivity effective mass, E_i is the Fermi level energy of the intrinsic semiconductor, which generally lies very close to the middle of the bandgap. It is clear that only the bulk traps whose energy are near the middle of the bandgap are effective recombination. Considering only these traps, Eqn. 2.12 can be reduced to 2.14.

$$U_t = \frac{\sigma_n \sigma_p v_{th} N_t (pn - n_i^2)}{\sigma_n [n + n_i \exp(\frac{E_t - E_i}{kT})] + \sigma_p [p + n_i \exp(\frac{E_t - E_i}{kT})]} \tag{2.12}$$

$$v_{th} = \sqrt{\frac{3kT}{m_c}} \tag{2.13}$$

$$U_t = \frac{\sigma_n \sigma_p v_{th} N_t (pn - n_i^2)}{\sigma_n (n + n_i) + \sigma_p (p + n_i)} \tag{2.14}$$

For low-level injection in n-type Si, the net bulk trap transition rate can be simplified as shown in Eqn. 2.15. The carrier lifetime due to bulk trap recombination is defined by Eqn. 2.16. In the case of high-level injection, Eqn. 2.14 can be further reduced to 2.17 and the excess carrier lifetime is then defined by Eqn. 2.18. It is clear that the lifetime due to trap recombination increases with the injection level.

$$U_t = \frac{\sigma_n \sigma_p v_{th} N_t [(p_{n0} + \Delta p)n - n_i^2]}{\sigma_n n} \approx \sigma_p v_{th} N_t \Delta p \tag{2.15}$$

$$\tau_{tp} = \frac{1}{\sigma_p v_{th} N_t} \tag{2.16}$$

$$U_t \approx \frac{\Delta n \sigma_n \sigma_p v_{th} N_t}{\sigma_n + \sigma_p} \tag{2.17}$$

$$\tau_{tHL} = \frac{\sigma_n + \sigma_p}{\sigma_n \sigma_p v_{th} N_t} = \tau_{tp} + \tau_{tn} \tag{2.18}$$

From the above discussion, it is obvious that the bulk trap recombination is the dominant mechanism when the injection level is low, while the band-to-band recombination is the dominant mechanism under high-level injection.

The excess carrier lifetime is one of the most important parameters for bipolar power devices. Increasing the lifetime will lead to a higher stored charge density in the drift region. This could cause the decrease of on-state voltage drop but with the sacrifice of switching speed. Hence control of lifetime is desired in the fabrication of devices for different applications. Two commonly used lifetime control methods are injection of gold and electron irradiation [24]. The energy level of gold atom trap is close to the middle of the bandgap. They become efficient recombination centers when injected into Si. Therefore, the lifetime decreases with the increasing gold doping density. In the case of high energy electron irradiation, electrons penetrate deeply into the Si and damage the lattice. The imperfections of crystalline lattice act as recombination centers and can be controlled by varying the irradiation level.

2.1.1.3 Drift and Diffusion

The current flows in a semiconductor is the sum of the net flow of holes in the direction of the current and the net flow of electrons in the opposite direction. The free carriers can move via two mechanisms, drift and diffusion.

When an electric field is applied across a semiconductor, the free carriers are accelerated by the field. Holes move toward the negative pole, while electrons move toward the positive pole. Such movements are called drift. The velocity with which they move is termed drift velocity and is proportional to the strength of the electric field $\mathscr{E}$, as shown in Eqn. 2.19. Where μ is the carrier mobility.

$$v_d = \mu \mathscr{E} \tag{2.19}$$

Carrier mobility is an important parameter for power semiconductor devices, since it directly affects the conductivity of the semiconductor. In Si, it is significantly affected by the carrier scattering caused by acoustic phonons and ionised impurities. The acoustic phonons are caused by lattice vibrations which decrease with decreasing temperature. The impurity scattering effect is caused by the interaction between a free carrier and the added impurities due to Coulomb scattering. The carrier mobility can be approximated according to Matthiessen rule, Eqn. 2.20, by combining the influences from different sources of scattering. Here μ_l is the mobility if only acoustic phonon scattering is present and μ_i is the mobility under the effect of ionised impurities.

$$\mu = (\frac{1}{\mu_l} + \frac{1}{\mu_i}) \tag{2.20}$$

The mobility decreases with the effective mass according to the equations in [24]. The current due to carrier drift under electric field is given by Eqn. 2.21. Where $\mathscr{E}$ is the applied electric field, μ_n is the electron mobility, and μ_p is the hole mobility.

$$J_{drift} = nq\mu_n\mathscr{E} + pq\mu_p\mathscr{E} \tag{2.21}$$

Diffusion is caused by the variation in spatial density of the free carriers. The nonuniform distribution of carriers may be caused by a variation in doping density or the excess carrier injection from a junction. The carriers tend to move from regions of higher concentration to regions of lower concentration due to its random thermal velocity. The flow of carriers is proportional to the concentration gradient and can be expressed by the Fick's law, as shown in Eqn. 2.22 which takes electrons as an example. The equivalent equations can be applied to holes as well. The current component due to diffusion hence can be given by Eqn. 2.23.

$$\frac{dn}{dt}|_x = -D_n\frac{dn}{dx} \tag{2.22}$$

$$J_{diffusion} = J_n + J_p = qD_n\frac{dn}{dx} - qD_p\frac{dp}{dx} \tag{2.23}$$

Where D_n and D_p are the electron and hole diffusion constant, respectively. The diffusion constants and mobilities are related by the Einstein relation, which is given by Eqn. 2.24. At room temperature (300 K), $kT/q = 0.026$ eV.

$$\frac{D_p}{\mu_p} = \frac{D_n}{\mu_n} = \frac{kT}{q} \tag{2.24}$$

Therefore, the total current density which consists of both drift and diffusion of free carriers can be given by Eqn. 2.25. Here E_{Fn} and E_{Fp} are quasi Fermi levels for electrons and holes, respectively.

$$J_{cond} = J_n + J_p = q\mu_n(n\mathscr{E} + \frac{kT}{q}\frac{dn}{dx}) + q\mu_p(p\mathscr{E} + \frac{kT}{q}\frac{dp}{dx}) = n\mu_n\frac{dE_{Fn}}{dx} + p\mu_p\frac{dE_{Fp}}{dx} \tag{2.25}$$

2.1.2 PN Junction

2.1.2.1 Thermal Equilibrium

The PN junction is the fundamental building block of nearly all power semiconductor devices. It is formed when a p-type region is adjacent to an n-type region. When a PN junction is formed, the majority carriers on either side of the junction will diffuse across it to the opposite side, where their density is much lower. The diffusion process will lead to the formation of space charge layers on both sides of the junction. Consequently an electric field is formed which retard the diffusion process by pushing the electrons back to n-type side and holes to p-type side. An equilibrium is reached when the effect of electric field counterbalanced with the diffusion and the total conduction current is zero. Hence Eqn. 2.25 can be expressed in the case of electrons, as shown in Eqn. 2.26. This means in thermal equilibrium condition, the quasi Fermi level must be constant throughout the sample.

$$J_n = n\mu_n \frac{dE_{Fn}}{dx} = 0 \tag{2.26}$$

Fig. 2.3 shows a step PN junction with $N_a > N_d$. The total negative charge in p-side must equal to the total positive charge in n-side according to thermal equilibrium condition, as shown in Eqn. 2.27. Since the doping concentration in p-type region is higher than that in n-type region, the depletion layer width in n-type region W_{Dn} must be longer than that in p-type region W_{Dp}.

$$N_a W_{Dp} = N_d W_{Dn} \tag{2.27}$$

Poisson's equation expresses the potential with the distribution of charge density. Eqn. 2.28 can be obtained by applying the Poisson's equation in the PN junction and assuming the charge density only varies one dimensional. Here, ε is the dielectric constant of Si (1.04×10^{-12} F/cm), $\rho(x)$ is the total charge density distribution along x coordinate.

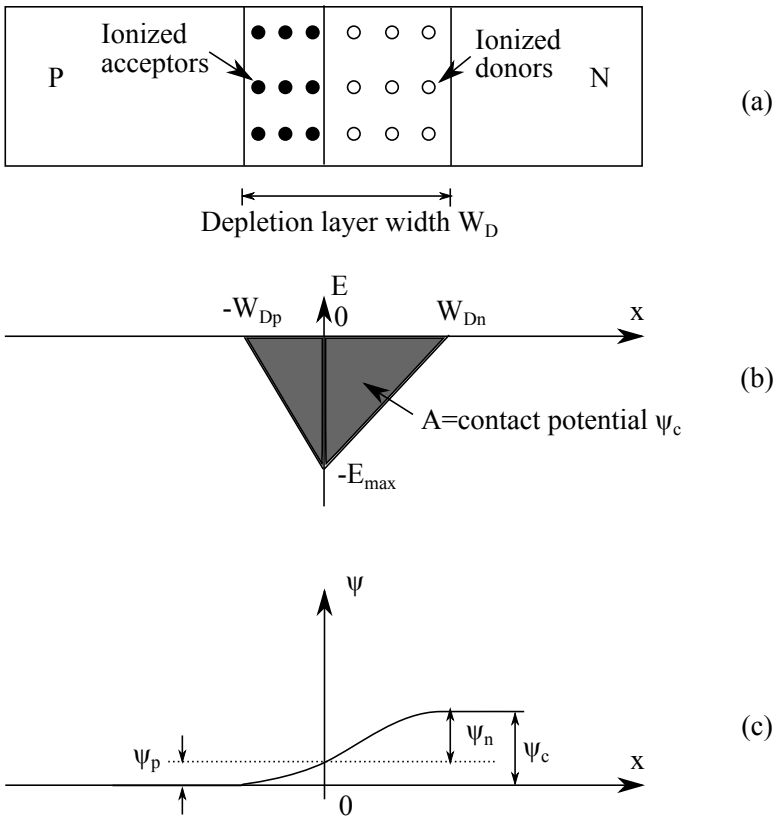


Figure 2.3: PN junction in thermal equilibrium. (a) Space charge distribution. (b) Electric field distribution. (c) Potential distribution.

$$-\frac{d^2\psi}{dx^2} = \frac{d\mathscr{E}}{dx} = \frac{\rho(x)}{\varepsilon} = \frac{q}{\varepsilon}[N_d^+(x) - n(x) - N_a^-(x) + p(x)] \tag{2.28}$$

Assuming complete ionisation in the depletion layer and $n(x) \approx p(x) \approx 0$. Eqns. 2.29 and 2.30 can be deduced for the depletion layer in p-side and n-side, respectively.

$$\frac{d^2\psi}{dx^2} = \frac{d\mathscr{E}}{dx} \approx \frac{qN_a}{\varepsilon} \tag{2.29}$$

$$-\frac{d^2\psi}{dx^2} = \frac{d\mathscr{E}}{dx} \approx \frac{qN_d}{\varepsilon} \tag{2.30}$$

Integrating Eqns. 2.29 and 2.30, the distribution of electric field can be obtained, as shown in Eqns. 2.31 and 2.32. As shown in Fig. 2.3, the maximum electric field appears at $x = 0$ and is given by Eqn. 2.33. The potential distribution can be obtained by integrating Eqns. 2.31 and 2.32, as shown in Eqns. 2.34 and 2.35. Therefore, the potentials across n-side and p-side can be found as Eqns. 2.36 and 2.37.

$$\mathscr{E}(x) = -\frac{qN_a(W_{Dp} + x)}{\varepsilon} \tag{2.31}$$

$$\mathscr{E}(x) = -\frac{qN_d(W_{Dn} - x)}{\varepsilon} \tag{2.32}$$

$$\mathscr{E}_{max} = \frac{qN_dW_{Dn}}{\varepsilon} = \frac{qN_aW_{Dp}}{\varepsilon} \tag{2.33}$$

$$\psi(x) = \frac{qN_a}{2\varepsilon}(W_{Dp} + x)^2 \tag{2.34}$$

$$\psi(x) = \psi(0) + \frac{qN_d}{2\varepsilon}(W_{Dn} - \frac{x}{2})x \tag{2.35}$$

$$\psi_p = \psi_c = \frac{qN_aW_{Dp}^2}{2\varepsilon} \tag{2.36}$$

$$\psi_n = \psi_c = \frac{qN_dW_{Dn}^2}{2\varepsilon} \tag{2.37}$$

The contact potential could also be expressed as the area of the triangle shown in Fig. 2.3b. Therefore, the depletion layer width can be calculated by Eqn. 2.38. In the case of the PN junction with $N_a \gg N_d$, the depletion layer width can be further reduced to Eqn. 2.39.

$$W_D = \frac{2\psi_c}{\mathscr{E}_{max}} = \sqrt{\frac{2\psi_c(N_a + N_d)\varepsilon}{qN_aN_d}} \tag{2.38}$$

$$W_D = \sqrt{\frac{2\psi_c\varepsilon}{qN_d}} \tag{2.39}$$

The contact potential caused by the electric field is given by Eqn. 2.40. In a silicon PN juction with $N_a = N_d = 10^{16}\ cm^{-3}$, the contact potential $\psi_c = 0.72$ eV at room temperature.

$$\psi_c = \frac{kT}{q} \ln \frac{N_aN_d}{n_i^2} \tag{2.40}$$

2.1.2.2 Forward and Reverse Bias

When an external voltage is applied across the PN junction, it appears mostly over the depletion layer since the free carrier density in this region is much lower than the rest parts, which means higher resistance. If the p-side is connected to the positive, the applied potential opposes the contact potential hence reduces the width of the depletion layer. This is termed forward biased PN junction, as shown in Fig. 2.4a. In the opposite case, the applied voltage increase the contact potential and lead to the expansion of the depletion layer. Such junction is named reverse biased PN junction, as shown in Fig. 2.4b.

The width of the depletion layer under applied voltage V (negative if the junction is

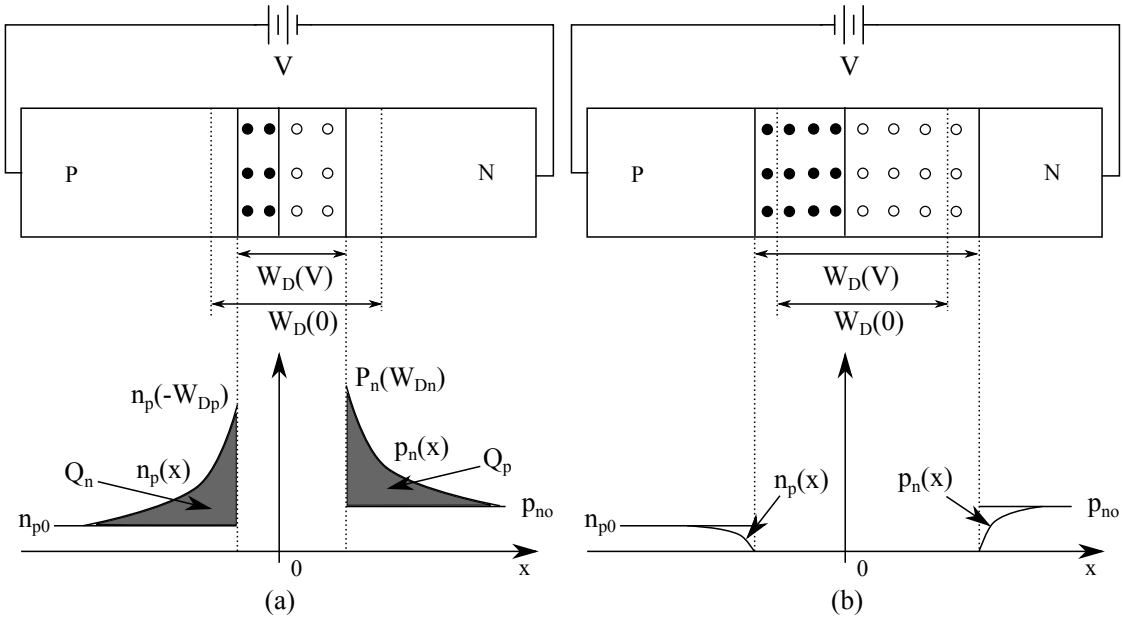


Figure 2.4: PN junction and its minority carrier densities in each side when it is (a) forward biased and (b) reverse biased.

reverse biased) is given by Eqn. 2.41. Here $W_D(0)$ is the depletion layer width under thermal equilibrium, which was expressed in Eqn. 2.38. The maximum electric field under this situation is given by Eqn. 2.42.

$$W_D(V) = W_D(0)\sqrt{1 - \frac{V}{\psi_c}} \tag{2.41}$$

$$\mathscr{E}_{max}(V) = \frac{2\psi_c}{W_D(0)}\sqrt{1 - \frac{V}{\psi_c}} \tag{2.42}$$

In the case of reverse bias, the minority carrier density at the edge of depletion layer become close to zero since any minority carriers appears in the depletion layer will be pushed across the PN junction by the electric field. The minority carriers in the rest of the material tends to diffuse toward the depletion layer due to the density gradients. Therefore a small leakage current is formed and is termed reverse saturation current (I_s). This current is highly

temperature sensitive since the minority carrier density in the rest part of the material increase significantly with temperature, as expressed by Eqns. 2.6 and 2.2 in the case of holes in n-type Si. Therefore, I_s needs to be considered when devices need to operate at high temperatures.

It is clear that a forward bias voltage lowers the potential barrier and the carriers are free to diffuse from the higher density side to the lower density side, for example the electrons will diffusion from n-side to p-side, as shown in Fig. 2.4a. It leads to a huge increase in the minority carrier densities adjacent to the depletion layer on both sides of the PN junction. This mechanism is called excess carrier injection and affects the characteristics of bipolar devices significantly, especially the switching behavior. The more excess carriers injected into the n- drift region, the lower on-state voltage and the longer turn-off duration.

The injected minority carriers recombine with the majority carriers as they diffuse from the PN junction, which causes its density to decrease exponentially with its diffusion distance, as illustrated in Fig 2.4a. The diffusion length is given by Eqns. 2.43 and 2.44 for electrons in p-side and holes in n-side, respectively.

$$L_n = \sqrt{D_n \tau_n} \tag{2.43}$$

$$L_p = \sqrt{D_p \tau_p} \tag{2.44}$$

The minority carrier density at the edge of the depletion layer, $n_p(-W_{Dp})$ and $p_n(W_{Dn})$ as shown in Fig. 2.4a, is given by Eqns. 2.45 and 2.46, respectively. Here, n_{p0} and p_{n0} are the thermal equilibrium electrons in p-type Si and holes in n-type Si, respectively.

$$n_p(-W_{Dp}) = \frac{n_i^2}{N_a} \exp(\frac{qV}{kT}) \approx n_{p0} \exp(\frac{qV}{kT}) \tag{2.45}$$

$$p_n(-W_{Dn}) = \frac{n_i^2}{N_d} \exp(\frac{qV}{kT}) \approx p_{n0} \exp(\frac{qV}{kT}) \tag{2.46}$$

When the forward bias reach its steady-state, the excess carriers recombination rate must be the same as the rate they injected across the junction, known as the forward bias current density (J). This equilibrium can be expressed by Eqn. 2.47. Where Q_n and Q_p are the total excess carrier charge in p-side and n-side respectively, as shown in Fig. 2.4a. They can be integrated from the known edge conditions given in Eqns. 2.45 and 2.46, since the excess carrier density decreases exponentially with the diffusion distance, as shown in Eqns. 2.48 and 2.49 [24].

$$J = \frac{Q_n}{\tau_n} + \frac{Q_p}{\tau_p} \tag{2.47}$$

$$Q_n = qn_{p0}[\exp(\frac{qV}{kT}) - 1]L_n \tag{2.48}$$

$$Q_p = qp_{n0}[\exp(\frac{qV}{kT}) - 1]L_p \tag{2.49}$$

The excess carrier density rather than the minority carrier density is used since only excess carriers are contributing to the diffusion current. Substitute Eqns. 2.48 and 2.49 into 2.47, the forward bias current density can be expressed as 2.50.

$$J = q(\frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n})[\exp(\frac{qV}{kT}) - 1] \tag{2.50}$$

From Eqn. 2.50, it is clear that when the PN junction is forward biased, the current density increases exponentially with the increasing voltage. While the PN junction is reverse biased, the current density is not zero, as discussed in the beginning of this section. The reverse saturation current can be expressed by Eqn. 2.51. From this equation, it is clear that I_s is almost independent of the reverse voltage. It is strongly temperature dependent since both p_{n0} and n_{p0} depend on the doping concentration and intrinsic carrier density n_i, which increases with temperature as discussed before.

$$I_s = q(\frac{D_p p_{n0}}{L_p} + \frac{D_n n_{p0}}{L_n}) \tag{2.51}$$

As the reverse voltage applied across the PN junction rises, the electric field will increase and accelerate the free carriers to a higher velocity. The kinetic energy of the free carrier increases squarely with the increase of its velocity. When its kinetic energy is high enough to break the covalent bond of the lattice, more free carriers will be formed and this process is termed impact ionisation. The newly liberated carriers can gain enough energy from the applied field and break another covalent bond. Therefore a large number of free carriers, i.e. large current, can be produced very quickly. This mechanism is called avalanche breakdown. During the breakdown process, the reverse voltage is nearly constant while the current increases dramatically, hence large amount of power is dissipated by the PN junction and will destroy the device if the reverse voltage is not reduced immediately.

The electric field required for the avalanche breakdown is termed $\mathscr{E}_{BD}$ and can be approximately estimated by a simple model. Assuming all the kinetic energy free carriers gain from the electric field are transferred to break the covalent bond, which minimum requirement is the bandgap (E_g). Define the time between collision is t_c, Eqn. 2.52 holds according to the law of conservation of energy. Here, the term qE_g is used as the energy required to break the band because E_g normally given in the unit of eV. The value of $\mathscr{E}_{BD}$ can be derived from this energy equation and is shown in Eqn. 2.53. The typical breakdown electric field for power devices measured in experiment is 200,000 V/cm. Its a very important parameter that limits the reverse blocking capability of power devices.

$$qE_g = \frac{1}{2}mv^2 = \frac{1}{2}m(\frac{\mathscr{E}_{BD} q t_c}{m})^2 = \frac{\mathscr{E}_{BD}^2 q^2 t_c^2}{2m} \tag{2.52}$$

$$\mathscr{E}_{BD} = \sqrt{\frac{2E_g m}{q t_c^2}} \tag{2.53}$$

2.1.3 Power Diode Structure and Operation

Diodes used in high power applications are more complicated in structure compared with those designed for low power applications, since they are designed to have high voltage blocking capability and low conduction loss. A common modification is to add a lightly doped N- epitaxial layer between the heavily doped n-type and p-type Si. This section will introduce the principles of this modification as well as the operation of power diode in detail.

2.1.3.1 Basic Structure and Principles of Power Diode

Fig. 2.5 shows the typical structure for a p-i-n diode. A lightly doped N- layer is grown on the top of a heavily doped N+ substrate. The typical doping concentration for each layer is shown on the figure. In high power applications, the requirement for reverse blocking voltage is normally very high. To avoid the avalanche break down, the electric field in the depletion layer needs to be kept below $\mathscr{E}_{BD}$. According to Eqns. 2.40 and 2.42, the maximum electric field will decrease if the doping concentration decreases. Therefore, the lightly doped drift region is designed to absorb the depletion layer of the reverse biased P^+N^- junction. Considering that the breakdown voltage V_{BD} is much higher than contact potential ψ_c, it can be estimated from Eqns. 2.39 and 2.42, as shown in 2.54.

$$V_{BD} \approx \frac{\varepsilon \mathscr{E}_{BD}^2}{2qN_d} \tag{2.54}$$

The depletion layer width is also highly dependent on the doping concentration according to Eqn. 2.39. Therefore, the most of the depletion layer contains in the lightly doped drift region. This parameter is very important for the design of non-punch-through (NPT) devices, since the width of drift region must be longer. The reverse blocking voltage is considered as the dark area shown in Fig. 2.6. Therefore, the minimum drift region width W_d required can be calculated and is shown in 2.55.

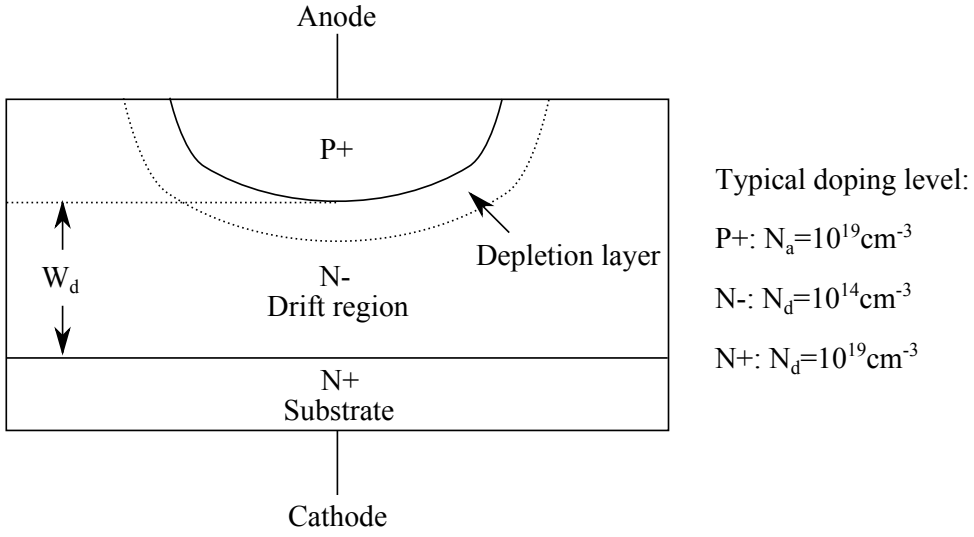


Figure 2.5: Cross section view of a typical p-i-n diode. W_d indicates the drift region width.

$$W_d \geq W(V_{BD}) = \frac{2V_{BD}}{\mathscr{E}_{BD}} \tag{2.55}$$

For power diode used in high voltage applications, the drift region width for NPT devices needs to be relatively long in order to accommodate the depletion layer, hence more Si is needed for the production of such device. Punch-through (PT) devices are designed to overcome this problem. Fig. 2.6 shows the distribution of depletion layer in NPT and PT devices. The same blocking voltage capability can be obtained with shorter drift region. In the case of PT devices, the increase of reverse voltage can not extend beyond the N- region thickness (W_d). The electric field at the edge of N-N+ junction starts to rise. When the doping concentration of the N- drift region is very low, the shape of the electric field is flatten out and becomes more rectangular, as shown in Fig. 2.6b. It is clear that the required drift region width for the PT devices are approximately one-half of that required for the NPT devices.

According to Eqn. 2.41, the depletion layer width of the drift region will be quite large because of the low doping level. This would appear to lead to high on-state resistance hence

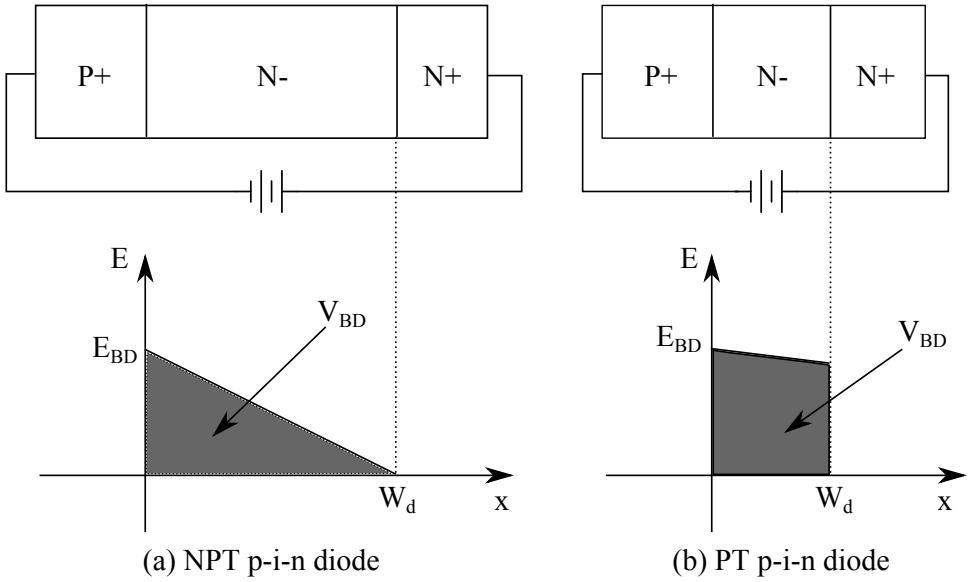


Figure 2.6: The distribution of depletion layer in (a) NPT and (b) PT p-i-n diode.

limit the forward conduction capability of power diode. However as a bipolar device, a large amount of excess holes are injected into the drift region from the forward biased PN junction during on-state. When the injected hole density (Δp) is greater than the thermal equilibrium electrons of the drift region n_{n0}, it cannot be neutralised by the electrons of the drift region. The positive space charge attracts the electrons to be inject into the drift region from the N+ region. When $\Delta p \gg n_{n0}$, the excess hole and electron densities are approximately equal. This double injection process is termed conductivity modulation. Such a mechanism reduces the on-state resistance of the drift region hence ensures the design of N- drift region practically valuable.

Fig. 2.7 shows the spatial distribution of the excess carrier caused by conductivity modulation. In the drift region, it is flattened if the diffusion length L is greater than the W_d. The current in the drift region can be written approximately as Eqn. 2.56 by assuming the electric field distribution in the drift region is uniform and the voltage across it is not high enough to cause carrier velocity to saturate. Here, V_d is the voltage over the drift region, A is the cross section area of the drift region, n_a is the excess carrier density defined in the

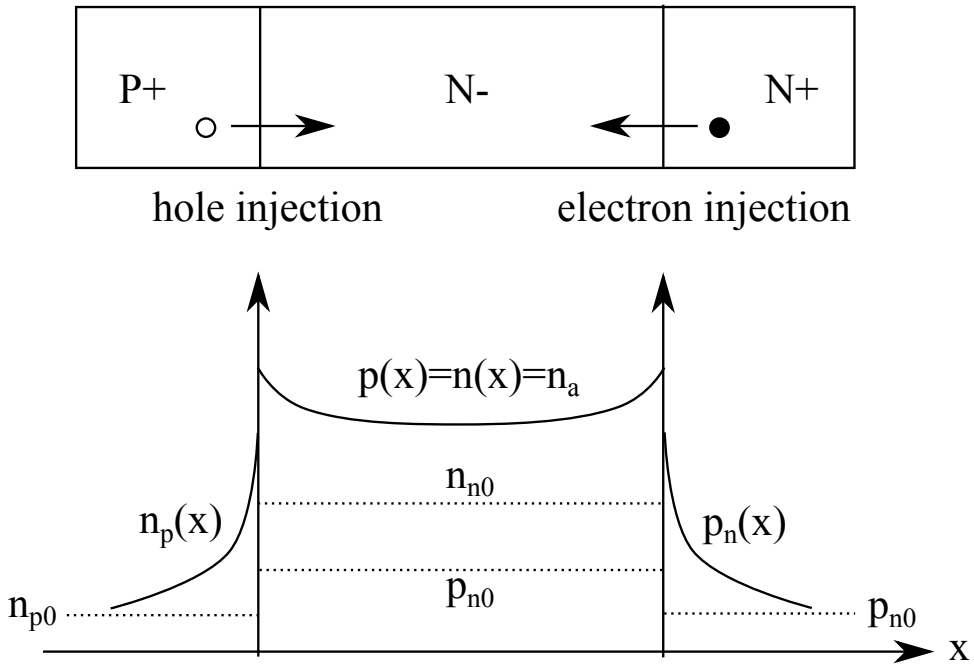


Figure 2.7: Conductivity modulation.

figure, and Q_d is the stored charge in the drift region. Therefore, V_d can be derived and is shown, Eqn. 2.57. The on-state voltage drop V_F across the diode is the sum of the voltage drop over PN junction V_j and V_d. V_j can be estimated according to Eqn. 2.50, hence V_F can be summarized as Eqn. 2.58. J_s, τ and μ are highly temperature dependent values as discussed before, hence V_F varies with temperature.

$$I_F \approx q\mu n_a A\mathscr{E}_d = \frac{q\mu n_a A V_d}{W_d} \approx \frac{Q_d}{\tau} \approx \frac{q n_a A W_d}{\tau} \tag{2.56}$$

$$V_d \approx \frac{W_d^2}{\mu\tau} \tag{2.57}$$

$$V_F = V_j + V_d \approx \frac{kT}{q} ln(\frac{J_F}{J_s}) + \frac{W_d^2}{\mu\tau} \tag{2.58}$$

Although conductivity modulation reduces the resistance of the drift region, V_F is still

limited by W_d and the current density. Therefore, the larger reverse blocking capability, the greater on-state power loss. In the case of high-level injection, the excess carrier density gets larger than the doping density, the band-to-band recombination in the drift region increase significantly and become dominant instead of the trap recombination. Therefore, the carrier lifetime decreases and lead to the increase of V_d. Furthermore, carrier mobilities decreases with the increasing excess carrier density. In the case of high-level injection when excess carrier densities reach a certain level, the Coulomb scattering between free carriers increases significantly hence limits the mobility.

2.1.3.2 Switching characteristics of Power Diode

The switching characteristics of power diode is affected by its intrinsic properties as well as the external circuit. Conductivity modulation reduces the diode on-state voltage drop by high-level carrier injection. However, the injected carriers limits the turn-off speed since they need to be extracted as the depletion layer build up. The current rate of change (di/dt) is controlled by the circuit stray inductance (L_S). The most important feature of diode switching characteristics is the turn-on voltage overshoot and reverse recovery current, they will be discussed in detail in this section.

Fig. 2.8 shows the representative power diode switching waveforms. The diode turn-on process can be divided into two parts, labeled as t_1 and t_2 in the figure, respectively. In the first phase, the space charge stored in the depletion region is removed from the edge of drift region towards to the middle as the forward current increases. The forward voltage across diode increases from negative to it's peak value V_{FP} as the depletion layer shrinks. This period ends when the depletion layer disappears and the PN junction becomes forward biased. During the second phase, the excess carriers are injected into the drift region from both ends, as shown in Fig. 2.9a. V_F starts to decrease as carrier density in the drift region rises, since the voltage drop across the drift region is reduced by conductivity modulation. This phase ends when the carrier distribution reaches the steady-state which is dependent

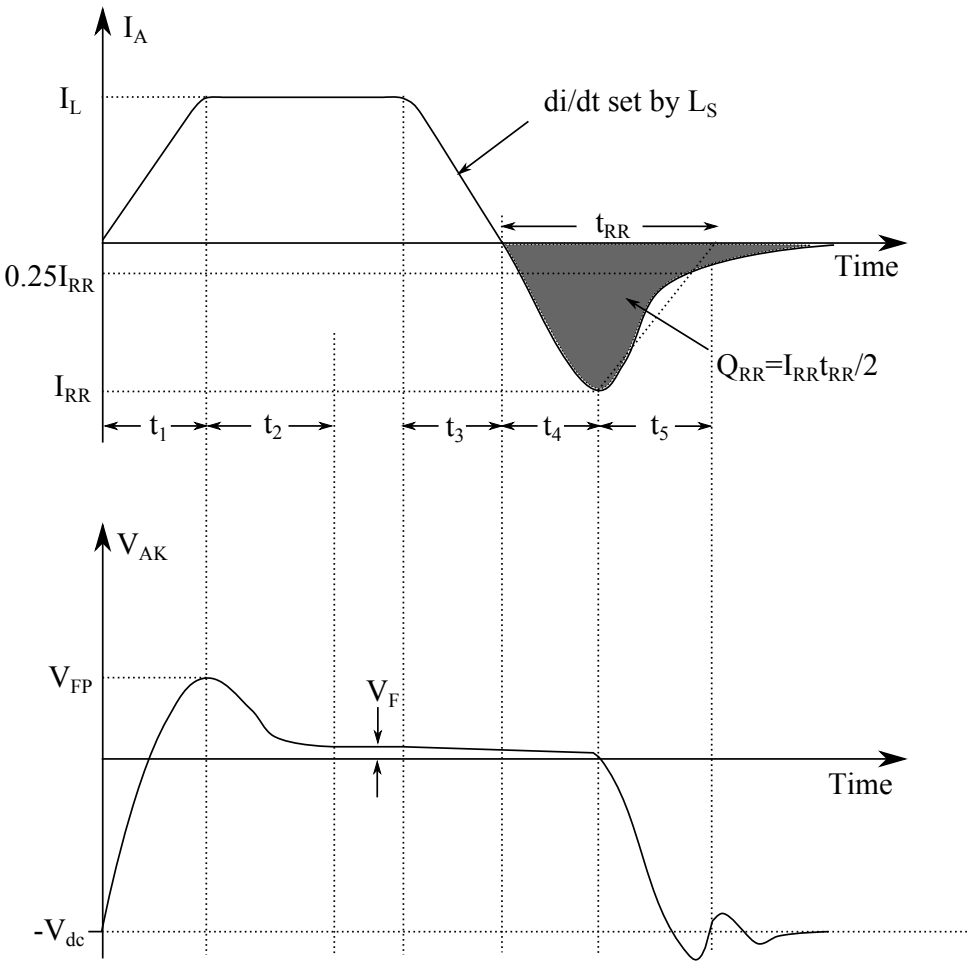


Figure 2.8: Power diode switching waveforms.

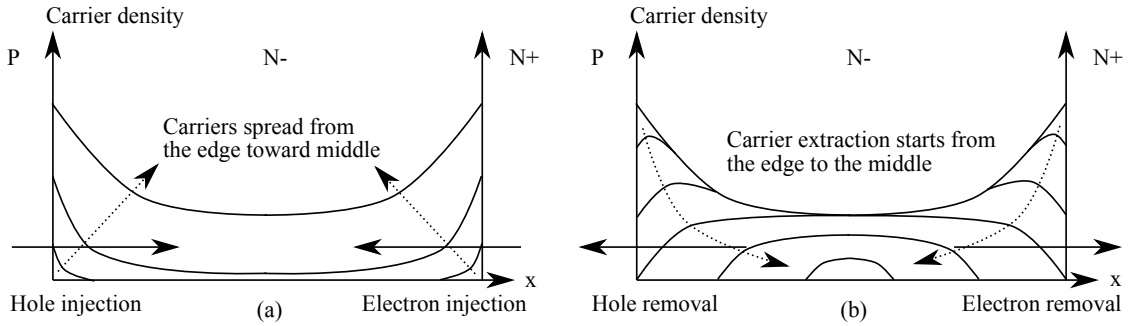


Figure 2.9: Carrier distribution in the drift region during (a) turn-on and (b) turn-off.

with forward current I_F. However, the turn-on voltage overshoot is not only caused by the high resistance of drift region when the excess carrier density is low. It is also affected by the diode series inductance L_d, which consists of the inductance of Si wafer and the bond wire. The voltage component that caused by L_d can be calculated by $L_d di/dt$.

The turn-on time depends on both di/dt and τ. A steeper current slope (di/dt) could shorten t_1 since less time is needed to remove the space charge. Also, a greater τ will shorten the time (t_2) needed for the excess carrier distribution to reach steady-state. However, a greater I_F tends to lengthen both t_1 and t_2 since di/dt is determined by external circuit and more time is needed to reach the steady-state.

The diode turn-off process consists of three phases. In the first phase which is labeled as t_3 in Fig. 2.8, the diode forward current begins to drop while V_F only drops slightly since the PN junction cannot be reverse biased until the excess carriers are removed from the drift region. During this period, the V_F drop is caused by the reduction of forward current. The second phase starts when the current goes negative. The current reduction rate di/dt, which is controlled by L_S, is maintained the same as the first phase by the excess carriers stored in the drift region. The second phase ends when the excess carrier density at the edge of the drift region reduces significantly and the PN junction becomes reverse biased. At this point, V_F turns negative and starts to build up rapidly towards $-V_{dc}$ as the depletion layer expands from the ends to the center of the drift region. In the third phase, the stored excess

carriers are not enough to support the negative current, hence the current starts to decay towards zero. V_{AK} will reach its steady-state value of $-V_{dc}$ after a few oscillations caused by the interaction between the device capacitance and circuit stray inductance.

The diode turn-off time depends on the time needed to extract the stored excess carriers in the drift region. Such time is governed by L_S which controls the speed of carrier sweepout di/dt, and the total excess carrier charge in the drift region (Q_F). Therefore, the turn-off time can be shortened by reducing Q_F with the price of increased on-state voltage drop. The stored excess carriers can also be removed by recombination. Therefore, shorter carrier lifetime will lead to faster switching, with the price again higher on-state voltage drop. The switching time can also be shorten by reduce W_d, which could cause a lower reverse blocking capability. These trade-offs must be considered to meet the requirements in different applications while designing power diodes.

2.1.4 IGBT Structure and Operation

The IGBT is a gate-controlled device which combines the advantages from both MOSFET and bipolar devices. Like the MOSFET, the IGBT is controlled by the gate voltage instead of current, which shortens its switch time and makes it possible to be used in high frequency applications. As a bipolar device, the on-state voltage drop of IGBT for high current applications is much lower than unipolar devices such as MOSFET due to conductivity modulation when their reverse blocking capability is the same. With these advantages, the IGBT can be found in a wide range of applications, for example electric vehicle and renewable energy. This section describes the basic structure, principles and switching characteristics of the IGBT.

2.1.4.1 Basic Structure and Principles of IGBT

Fig. 2.10 shows the basic structure of IGBT. It looks like a MOSFET with a p-type Si base added in the bottom. This structure appears that it could never conduct current from

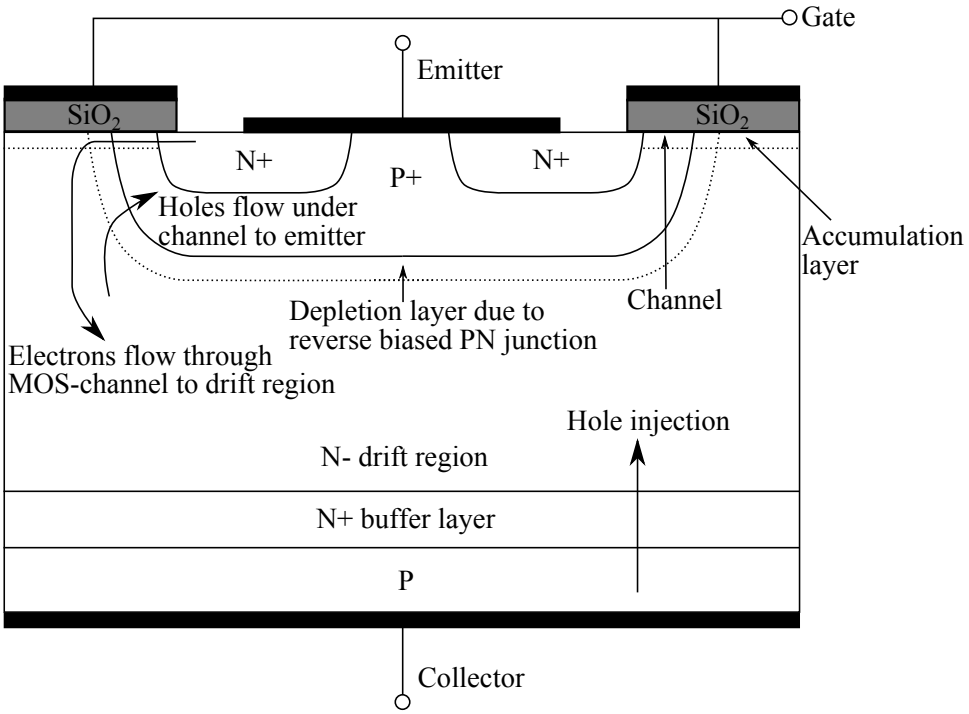


Figure 2.10: Basic IGBT structure and its behavior in the on-state.

collector to emitter since N-P+ junction will be reverse biased if the collector-emitter voltage V_{CE} is positive. However, the special gate design makes it possible to control its conductivity.

The gate terminal is isolated by a oxide layer and is placed on the top of the P-well. This structure can be thought of as a capacitor, termed gate oxide capacitance C_{OX}. When a small positive gate-emitter voltage V_{GE} is applied, the positive charge induced on the gate terminal requires the same amount of negative charge on the Si side. The electric field of the positive charge pushes the holes away from the Si-oxide interface hence forms a depletion layer. Further increase V_{GE}, the depletion layer starts to expand in thickness and starts to attract free electrons to provide the required negative charge. When V_{GE} increased to a certain level, the density of free electrons at the interface will become high enough to form a layer which is highly conductive, as shown in Fig. 2.10. Such value of V_{GE} is termed the threshold voltage V_{TH}. This layer is formed in a p-type Si but behaves as a n-type Si and conducting the IGBT on-state current, therefore it is usually called inversion layer or MOS

channel.

When $V_{GE} > V_{TH}$, the IGBT is turned on and able to forward conduct. The electrons flow through the MOS channel into the drift region and cause the same amount of holes to be injected from the collector into the drift region, as indicated in Fig. 2.10. This double injection will lead to conductivity modulation and reduce the on-state voltage drop (V_{CE}) just as it does in the diode. The injected holes move cross the drift region by both diffusion and drift. When they enters the P-well, their space charge will attract electrons from the emitter terminal hence the excess holes are quickly recombined. While the injected electrons move towards the opposite direction and enter the p-type collector region where they either recombine with the majority carriers or move into the collector metallisation.

When $V_{GE} < V_{TH}$, the charge of gate oxide capacitance is not high enough to support the MOS channel, therefore the IGBT is turned off and able to block the forward voltage. The principles of forward voltage blocking for IGBT is the same as that for power diode. The depletion layer of the reverse biased N-P+ junction expands according to V_{CE}. If the drift region is long enough to accommodate the depletion layer, the N+ buffer layer shown in Fig. 2.10 is no longer necessary. This type of IGBT is called NPT IGBT and able to block reverse voltage by the PN- junction. If the reverse blocking capability is not required, it is possible to reduce the thickness of the drift region by the PT structure as shown in the figure.

There are several parasitic components in the structure of IGBT, as shown in Fig. 2.11. These components greatly affect the behavior of IGBT. The gate-emitter capacitance (C_{GE}) which also refers to as the input capacitance, is the capacitance between gate and channel. This capacitance consists of C_{OX} and the capacitance of depletion layer forms at the Si-oxide interface, hence it varies with V_{GE}. The gate-collector capacitance (C_{GC}) which also refers to as the Miller capacitance, is the combination of C_{OX} and the capacitance of depletion layer in the drift region under the gate. This capacitance varies as the depletion layer area changes. during on-state, the depletion layer under gate disappears, hence the Miller capacitance

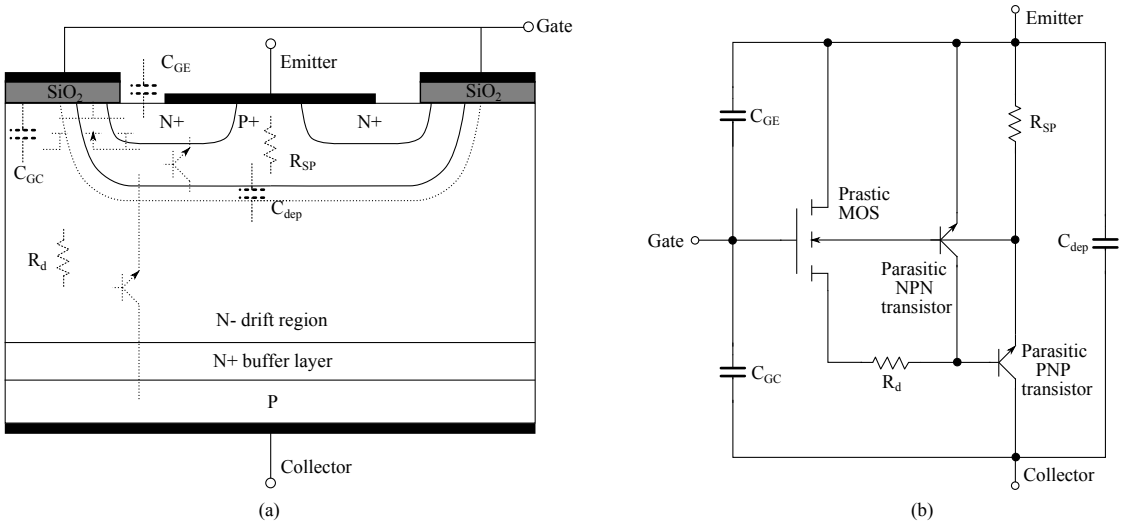


Figure 2.11: (a) IGBT structure with parasitic components and (b) accurate IGBT equivalent circuit.

approximately equal to C_{OX}. During off-state, C_{GC} is reduced since the depletion layer capacitance now connected in series with C_{OX}. Due to this feature, Miller capacitance could affect the switching characteristics of IGBT and lead to the gate voltage plateau which will be detailed in the next section. The collector-emitter capacitance C_{CE} is also referred to as the output capacitance. C_{CE} consists of both the depletion capacitance under the P-well layer C_{dep} and the capacitance of the charge stored in the drift region. Its value affects the rate of forward voltage change during switching dV_{CE}/dt.

Furthermore, parasitic PNP and NPN transistors can be found in the IGBT structure, as shown in Fig. 2.11. Together, they can form a parasitic thyristor. Under some specific conditions, this thyristor could be activated, which is called latchup and can cause the IGBT to become uncontrollable. During on-state, holes are injected from the P collector to the N- drift region and finally enters P-well where they recombine with the electrons from the emitter metallisation. There are two path ways of this hole flow, lateral and vertical across the P-well. Due to the negative charge of the inversion layer, most of the holes are attracted to flow laterally under it. A lateral voltage drop is then developed by this hole flow and the

resistance of the P-well (R_{SP}). Latchup occurs when this voltage is high enough to forward bias the P+N+ junction. Latchup could also occur during turn-off when dV/dt is too large. During turn-off, the excess carriers stored in the drift region is removed and depletion layer builds up. If dV/dt is too large, the discharging current will become large enough to forward bias the P+N+ junction and lead to the latchup of IGBT.

Obviously, latchup is not desired for IGBT operation, several modifications can be made to avoid it. In the device aspect, latchup can be avoid by reducing R_{SP}. This can be done by either reducing the lateral width of the N+ emitter region or increasing the doping concentration of the P-well. However, the doping concentration of the MOS channel region need to be kept relatively low to ensure the inversion layer could take place easily. In the circuit aspect, latchup can be avoid by slowing down the turn-off process to limit dV/dt. This could be accomplished by using a larger value of gate resistance (R_g), which limits the gate discharging current.

2.1.4.2 Switching characteristics of IGBT

Understanding the switching characteristics of IGBT is essential to estimate its switching loss. Fig. 2.12 shows the detailed switching waveforms of IGBT under inductive load condition. As shown in the figure, the switching process can be divided into ten phases, five for turn-on process and five for turn-off process.

The ten phases for IGBT switching process are listed below:

1. Gate voltage rise to V_{TH}.
2. I_C increase.
3. I_C recovery due to diode reverse recovery.
4. V_{GE} plateau.
5. V_{GE} increase to $V_{GG(on)}$.

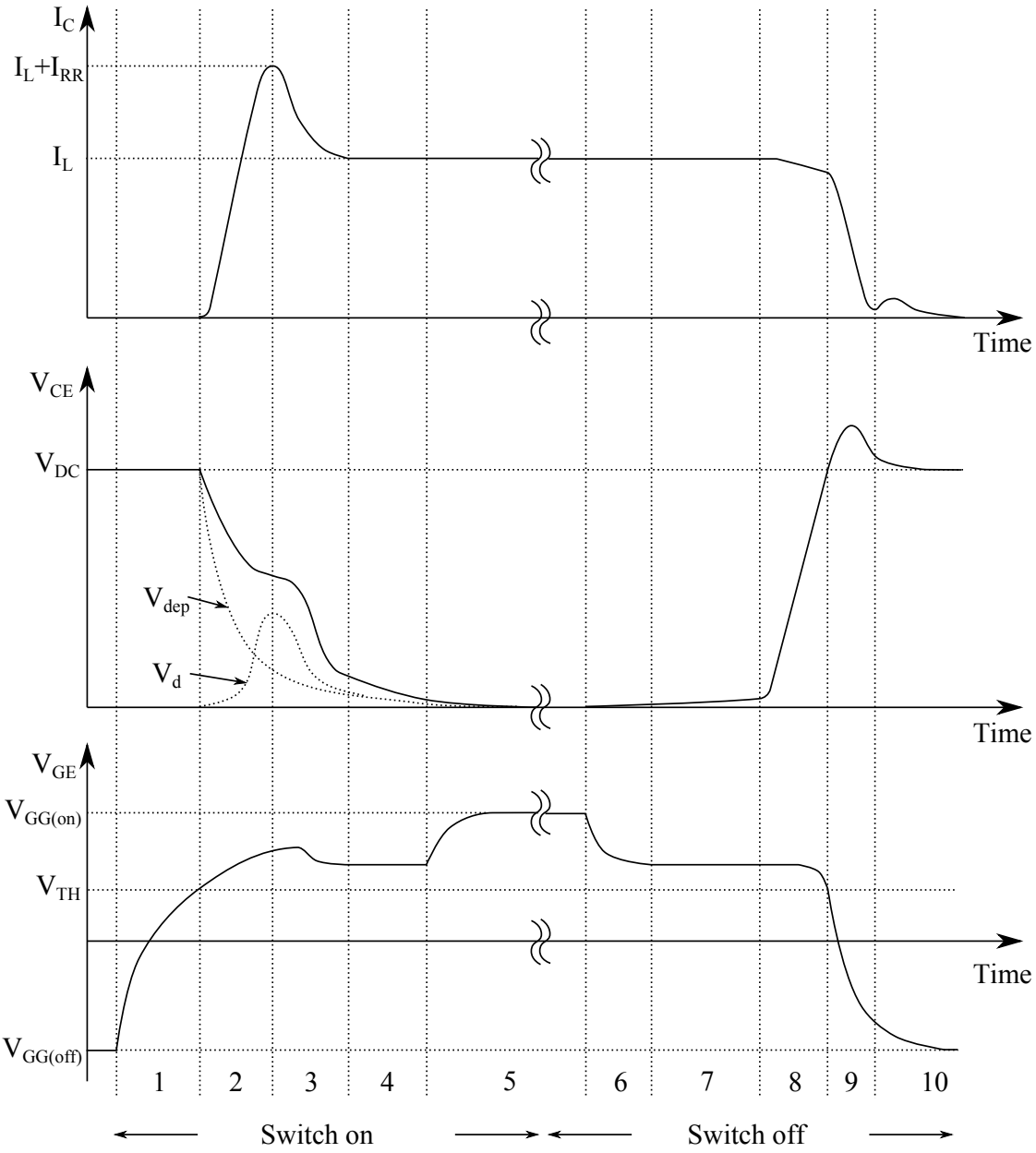


Figure 2.12: Detailed IGBT switching waveforms under inductive load condition.

6. V_{GE} fall.
7. V_{GE} plateau.
8. V_{CE} rise.
9. I_C fall.
10. I_C decay to zero.

The IGBT turn-on process starts from phase 1, when a positive gate drive voltage $V_{GG(on)}$ is applied between the gate and emitter. V_{GE} starts to rise towards V_{TH} and charges C_{GE} and C_{GC}. Therefore, the time constant of the V_{GE} rise curve can be given by $\tau_1 = R_g(C_{GE}+C_{GC})$. C_{GE} is small when IGBT is off as discussed in previous section, hence the rise of V_{GE} is dominate by C_{GC}. IGBT is off until V_{GE} increase above V_{TH} at the beginning of phase 2.

In phase 2, V_{GE} continues to rise as in phase 1. The inversion layer is built by V_{GE} and starts to conduct current, hence I_C begins to rise. As electrons inject from the emitter to the drift region from the MOS channel, holes starts to inject from the p-type collector into the drift region to neutralise the space charge. The excess carrier density starts to rise hence the depletion layer shrinks and lead to the drop of V_{CE}. V_{CE} in this phase is controlled by the combination of voltage drop across depletion layer (V_{dep}) and the drift region voltage drop (V_d). V_{dep} decreases exponentially as the depletion layer retreats while V_d increases as I_C rises. Since load current (I_L) is almost constant, I_C rises to its maximum value ($I_L + I_{RR}$) at the end of phase 2 due to the diode reverse recovery current. The MOS channel in this phase is saturated since V_{CE} is still high. Therefore, the MOS current (I_{MOS}) can be calculated by Eqn. 2.59. Here, K_{PL} is the MOS transconductance coefficient.

$$I_{MOS} = \frac{K_{PL}}{2}(V_{GE} - V_{TH})^2 \tag{2.59}$$

In phase 3, I_C drops from its maximum value towards its steady state value I_L as the diode current starts to decay towards zero. V_{dep} continues to decrease as the depletion layer

shrinks towards the P-well. V_d starts to drop from its maximum value with the decrease of I_C. The drop of these two voltages leads to the further decrease of V_{CE}. As V_{CE} decreases, the depletion layer under gate shrinks laterally from the MOS channel towards the center of interchip region. The positive gate charge started to attract free electrons under the gate and forms the accumulation layer, which lead to a significant increase of C_{GC}.

During phase 4, V_{CE} continues to drop because of the depletion layer shrink and the V_d drop due to increasing free carrier density in the drift region. However, the voltage across the MOS channel is still high enough to maintain its operation in the saturation region. V_{GE} is therefore clamped to the plateau value to main the constant I_C. Most of the positive gate current (I_g) flows into C_{GC} to help build up the accumulation layer. I_g can be calculated from Eqn. 2.60, with $V_{GE,I_{MOS}}$ indicates the required V_{GE} to support the MOS current. According to the equivalent circuit shown in Fig. 2.11, this could cause V_{CE} to decrease at a rate shown in Eqn. 2.61.

$$I_g = \frac{V_{GG(on)} - V_{GE,I_{MOS}}}{R_g} \tag{2.60}$$

$$\frac{dV_{CE}}{dt} = \frac{dV_{GC}}{dt} = \frac{I_g}{C_{GC}} \tag{2.61}$$

Once V_{CE} decreases significantly and close to its on-state value, the MOS channel enters the linear operation zone. V_{GE} become unclamped and starts to rise towards $V_{GG(on)}$ at a time constant of $\tau_2 = R_g(C_{GE} + C_{GC})$. Note that $\tau_2 > \tau_1$ since C_{GC} increases as the accumulation layer forms. The increase of V_{GE} will lead to the further reduction of MOS channel voltage hence V_{CE} drops slightly with the rise of V_{GE} and finally reach it steady on-state value.

The IGBT turn-off process is almost the inverse sequence of the turn-on process and consists of the last 5 phases. At the beginning of phase 6, the positive gate drive voltage ($V_{GG(on)}$) suddenly change to a negative value $V_{GG(off)}$. This causes the discharge of C_{GE}

and a negative gate current through R_g. MOS channel operates in the linear region during on-state, hence the MOS channel voltage increases with the drop of V_{GE}, which leads to the slight rise of V_{CE} and a constant I_C. The voltage drop of V_{GE} due to the discharging of gate capacitance follows the time constant τ_2.

At the start of phase 7, V_{CE} is high enough to force the MOS channel to enter the saturation region. Hence V_{GE} is clamped to a constant value to hold the constant load current and consequently enters the plateau region. Therefore the negative gate current is contributed by the discharging of Miller capacitance (C_{GC}). As the accumulation layer disappears from the center of the interchip area towards the MOS channel, C_{GC} starts to decrease and the resistance of the drift region increases which lead to the rise of V_{CE} and can be calculated by Eqns. 2.60 and 2.61.

The process enters phase 8 when C_{GC} is reduced significantly and its discharging current cannot support the negative gate current. C_{GE} starts to discharge hence lead to the decrease of V_{GE} and consequently I_C. Less excess carriers are injected into the drift region and the depletion layer, or V_{CE}, builds up quickly.

At the beginning of phase 9 as V_{GE} falls below V_{TH}, the MOS channel vanishes hence all the DC voltage are now applied over the IGBT. As the MOS channel disappears, I_C drops sharply and produces a voltage across the circuit stray inductance (L_S) in the same direction as V_{DC}. This leads to the V_{CE} turn-off overshoot and therefore forward bias the free-wheeling diode. The current then commutates from IGBT to diode. The V_{CE} overshoot decays when di/dt decreases as the current commutation between IGBT and diode completes. Note that, there might be some small oscillation in I_C at the end of this phase due to the interaction between the capacitance of IGBT and the stray inductance. V_{GE}, I_C and V_{CE} reaches their final stable value in phase 10, where the turn-off process terminates.

2.2 Physics Based IGBT and Diode Model

In this thesis, physics based IGBT and diode models are used to calculate the device switching power loss under different operating conditions. The desired model should be accurate, fast and temperature dependent. The current available device modelling techniques in literature are reviewed and compared. The temperature effects for the device operation are also discussed in the end.

2.2.1 Review of Device Modelling Technique

There are many papers on IGBT models since the invention of IGBT in 1980's. They can be categorised into three types according to their complexity, accuracy and mathematical method: behavioral models, finite element models and semi-numerical models. In the case of system-level simulation, the desired model should be accurate enough to calculate the power losses for both on-state and switching at different operating conditions. Also, they need to be fast enough for practical applications.

The behavioral models [28,29] simulate the device behavior without considering their internal physics. They are based on the measurements of device characteristics under different operating conditions. These empirical data are fitted with different methods to represent the IGBT behavior. Hence their accuracy depends on the density of measurement points. They are the fastest models since they ignore the device physics. However, the price for the speed is that they cannot give valid results when the device operates outside the fitted range. Furthermore, the internal device behavior such as carrier density and electric field distribution cannot be modeled. Hence this type of model is not suitable for the device design. The most applicable area for the behavioral models is the system-level modelling which requires fast simulation.

Accurate device behaviors can be obtained by solving all the physics based semiconductor equations without assumptions. These equations are highly nonlinear, some of them are

partial differential equations (PDEs). Therefore, finite element method is widely employed to solve these equations. The basic idea of finite element method is to find the approximate solutions for the PDEs using standard techniques such as Euler's method [30] or Runge-Kutta method [31]. Models using the finite element method [32] require accurate physical equations as well as necessary device data, such as device geometrical information, contact structures, doping densities and carrier lifetime, to solve the equations. The region under study is discretised on to a fixed mesh. The potential and carrier densities are defined at each mesh point. With known boundary condition, the electric field and currents can be derived by solving the PDEs. The results are then used to generate the potential and carrier densities of the neighboring points on the mesh. The accuracy of this type of model depends heavily on the density of the mesh. Normally, the mesh density of the area with steep changes in gradient, such as junctions, are higher than the other areas. This type of model is able to simulate the detailed carrier density and electric field distribution with a well defined mesh and accurate device data. Therefore, they are commonly employed in the device design and optimisation. However due to the time consuming numerical calculations, they are too slow to apply in circuit simulators (one switching event could take hours to run).

The semi-numerical models are partly based on device physics. Some simple models [33, 34] are developed with lumped-charge modelling technique which significantly simplify model equations without losing the basic structural information and internal physical processes in the devices. The basic idea of lumped-charge technique is to separate the space charge region into several specific regions. The carrier densities are assumed to be constant within each region. The electric field and current for each region are calculated based on the carrier density distribution by the physical equations. With these simplifications, the lumped-charge models are faster than the ambipolar diffusion equation (ADE) models, However, with less accuracy.

ADE solution models [35, 36] are more complex than the lumped-charge models. For a ADE model, the electric field and carrier storage distribution in the drift region are based on

the PDEs which is derived from device physics with simple boundary conditions while the other part of the device is described by simplified models. Therefore numerical calculations are needed to solve the PDEs. Different approaches for solving the ADE are reported. Strollo uses Laplace transform to transform the ADE from the time domain into the s-domain [37]. Morel et at. solve the ADE by internal approximation [38, 39]. Bryant and Palmer [36, 40] use Fourier series to solve the ADE and implement the device model in Matlab. Matlab is a widely used programming environment for algorithm development, data analysis and numerical computation. Therefore it is an ideal environment for system-level simulation. The implementation of device model in Matlab allows more compatibility in system simulation and data analysis hence is a desired feature. This model is able to achieve high accuracy with relatively short simulation run time (around 2 seconds for each switching event). The comparison between experiment switching waveforms and simulation results of this model is shown in a previous paper [22]. Furthermore, a simple parameter extraction method is offered for this model to make it practically applicable for different applications [41].

Although the simulation time needed for Bryant's model is already very short, it still cannot be applied to system-level simulation directly. However, it can be used in a circuit simulator to generate desired device parameters with one switching event under different operating conditions. In this work, device junction temperature is the only desired output from electrothermal simulation. Therefore, power loss LUTs can be generated by running the device model at different operating conditions. During the inverter electrothermal simulation, the inverter model generates the power loss by interpolating these LUTs. This could greatly increase the simulation speed while maintaining its accuracy.

2.2.2 Fundamentals of the ADE model based on Fourier series solutions

The behavior of high voltage devices are heavily affected by its lightly doped drift region. For example, the on-state voltage drop can be reduced by increasing the carrier lifetime in

the drift region. However, this would lead to higher turn-off loss since more free carriers are stored in the drift region during the on-state. The carrier lifetime is affected by its density especially in the case of high-level injection. When the excess carrier density exceeds a certain level, the Coulomb scattering between free carriers becomes serious and reduces their lifetime. The current and hence electric field distribution is related with the carrier density distribution as well. Therefore, the ability to simulate the excess carrier density is critical for a physics based device model.

Ambipolar diffusion is the simultaneous diffusion of positive and negative particles due to their interaction with electric field [42]. Fig. 2.13 shows the general arrangement of the depletion layers and carrier storage region (CSR) in the drift region. x_1 and x_2 are free to move according to different conditions. Assuming the change of carrier density in the drift region is one dimensional, ADE can be used to describe the behavior of free carriers in the drift region under high-level injection, as shown in Eqn. 2.62. Here D is the ambipolar diffusion constant, $p(x,t)$ is the carrier dynamic distribution function, and x is defined as shown in the figure.

$$D\frac{\partial^2 p(x,t)}{\partial x^2} = \frac{p(x,t)}{\tau_{HL}} + \frac{\partial p(x,t)}{\partial t} \tag{2.62}$$

Fourier series can be used to decompose periodic functions into the sum of a set of simple sine and cosine functions. In this case, $p(x,t)$ can be expressed by a Fourier series as shown in Eqn. 2.63.

$$p(x,t) = p_0(t) + \sum_{k=1}^{\infty} p_k(t) cos(\frac{\pi k(x-x_1)}{x_2-x_1}), \tag{2.63}$$

where,

$$p_0(t) = \frac{1}{x_2-x_1}\int_{x_1}^{x_2} p(x,t)dx \tag{2.64}$$

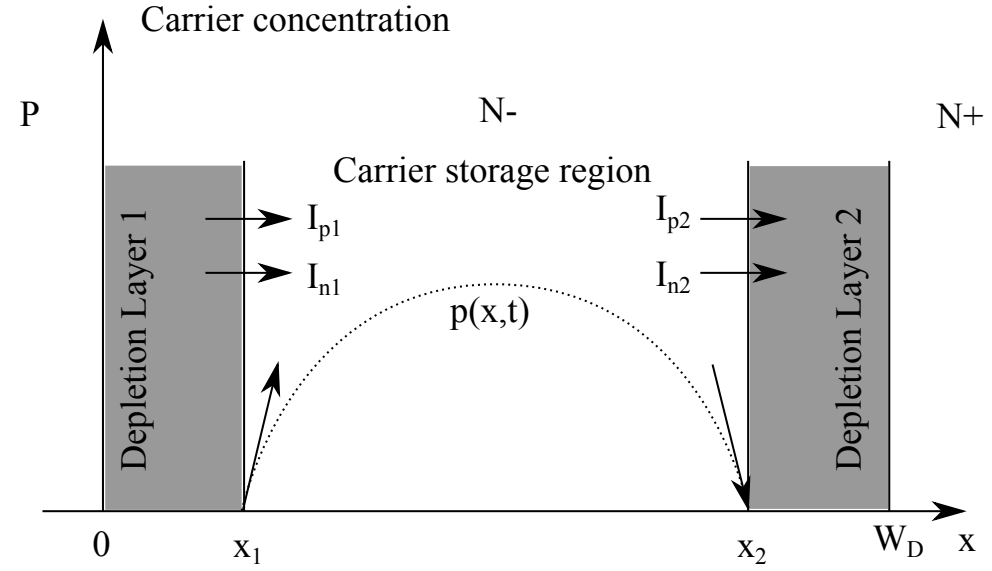


Figure 2.13: General arrangement of the depletion layers and CSR in the drift region.

$$p_k(t) = \frac{2}{x_2 - x_1} \int_{x_1}^{x_2} p(x,t) cos(\frac{\pi k(x - x_1)}{x_2 - x_1}) dx \tag{2.65}$$

This expression uses the cosine part of the Fourier series. Hence the odd term represents the uneven carrier density distribution profile while the even term represents the even carrier density distribution profile, as shown in Fig. 2.14.

The detailed derivation of the Fourier solution for the ADE is given in [43]. Its results are shown in Eqns. 2.66 and 2.67.

$$\frac{dp_0}{dt} = \frac{D}{x_2 - x_1}[\frac{\partial p}{\partial x}|_{x_2} - \frac{\partial p}{\partial x}|_{x_1}] - \frac{p_0}{\tau} - \frac{1}{x_2 - x_1} \sum_{n=1}^{\infty} p_n [\frac{dx_1}{dt} - (-1)^n \frac{dx_2}{dt}] \tag{2.66}$$

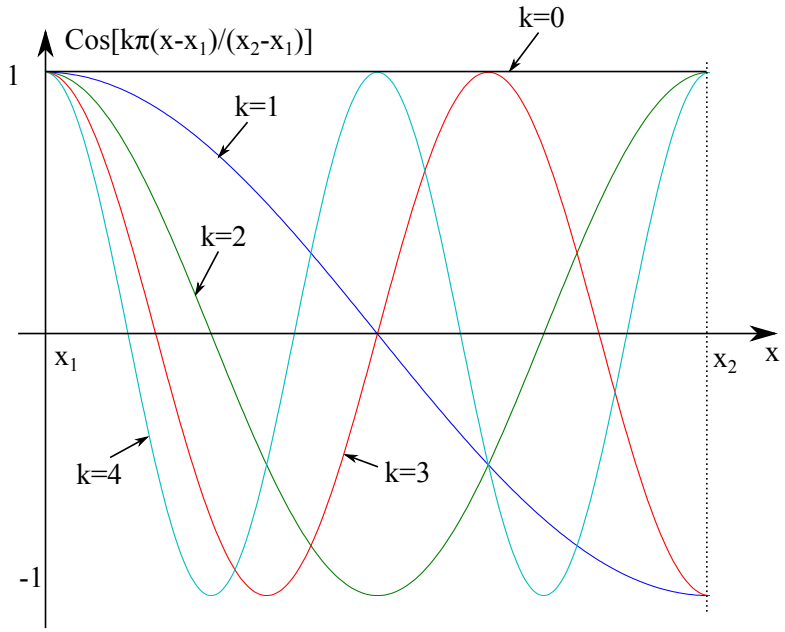


Figure 2.14: Carrier density profiles for odd and even term in the Fourier series.

$$\begin{aligned}\frac{dp_k}{dt} &= \frac{2D}{x_2-x_1}\left(\frac{\partial p}{\partial x}|_{x_2}(-1)^{-k}-\frac{\partial p}{\partial x}|_{x_1}\right)-p_k\left(\frac{1}{\tau}+\frac{D\pi^2k^2}{(x_2-x_1)^2}\right)\\ &+\frac{2}{x_2-x_1}\left(\sum_{n=1,n\neq k}^{\infty}\frac{n^2p_n}{n^2-k^2}\left(\frac{dx_1}{dt}-(-1)^{n+k}\frac{dx_2}{dt}\right)+\frac{p_k}{4}\left(\frac{dx_1}{dt}-\frac{dx_2}{dt}\right)\right)\end{aligned} \quad (2.67)$$

Eqns. 2.66 and 2.67 require the boundary conditions at the edge of the CSR. They are the position of the boundary (x_1, x_2) and the carrier density gradients ($\partial p/\partial x|_{x_1}$, $\partial p/\partial x|_{x_2}$). The carrier density gradient can be obtained from the boundary currents by rearranging the current transport equations. The drift current is given by Eqn. 2.21 while the diffusion current is given by Eqn. 2.23. Assuming $p = n$ and consider Eqn. 2.24, the carrier density gradient is given by Eqn. 2.68.

$$\frac{\partial p}{\partial x} = \frac{1}{2qA}\left(\frac{I_n}{D_n}-\frac{I_p}{D_p}\right) \quad (2.68)$$

The boundary currents can be obtained from the device terminal currents and the emitter recombination. The voltage drop across the depletion layer is calculated from the boundary carrier densities according to Eqn. 2.28. The depletion layer width can be calculated from the voltage drop according to Eqns. 2.39 and 2.41. The position x_1 and x_2 are then derived according to the depletion layer behavior. The other parts of the device model is based on the current transport equations or some previous developed model. The detail of the diode and IGBT model used in this thesis is presented by Bryant in [43] and will not be discussed here.

2.2.3 Temperature Effects

As discussed in previous sections, some properties of Si are highly temperature dependent. They are the bandgap, intrinsic carrier densities, carrier mobilities and carrier lifetime, which greatly affect the device characteristics. Therefore, the relationship between temperature and these material properties need to be considered and included in the device model. Furthermore, two other IGBT parameters are affected by temperature. They are threshold voltage and transconductance coefficient. The temperature effects on the device characteristics due to these parameter variation are summarised and discussed in this section.

According to Eqn 2.1, E_g decreases with the increasing temperature. Higher temperature means more thermal energy stored in the Si lattice as crystal vibrations hence more electrons are freed to move around the crystal. The reduction in E_g will lead to a smaller $\mathscr{E}_{BD}$ and hence smaller V_{BD}, according to Eqn. 2.53. n_i depends on both E_g and temperature according to Eqn. 2.2. Combine Eqns. 2.1 and 2.2, the relationship between intrinsic carrier density and temperature can be plotted as shown in Fig. 2.15. It is clear that n_i rises significantly with the increasing temperature. The increase of n_i will affect several device properties. First of all, it will lead to higher device leakage current according to Eqn. 2.51. More free carriers will diffuse towards the edge of the depletion layers where the carrier density is nearly zero. Secondly, the contact potential decreases with the increasing

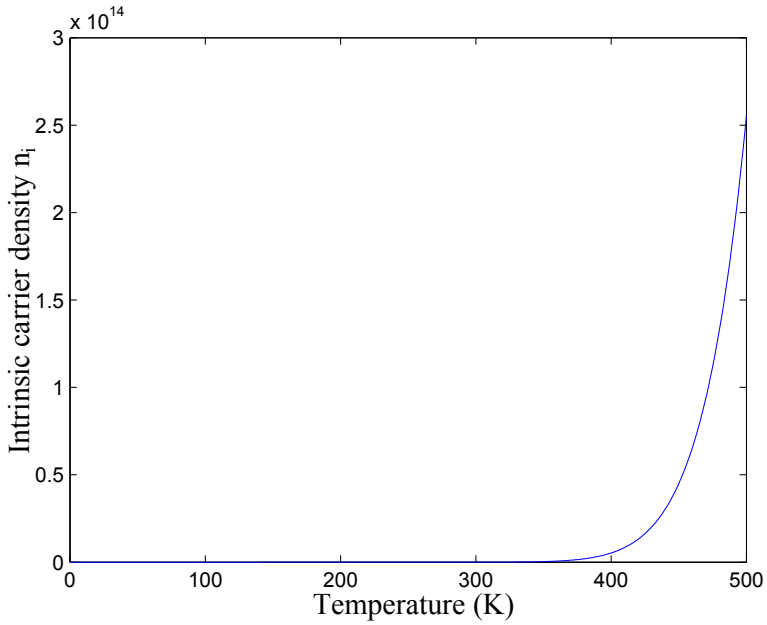


Figure 2.15: Intrinsic carrier density changes with temperature.

temperature according to Eqns. 2.2 and 2.40, as shown in Fig. 2.16. The exponentially increase of n_i limits the upper operation temperature of Si devices. When the temperature increases to a certain level (around 250 ℃ as shown in Fig. 2.15), the thermal energy inside the Si atom is high enough to break the covalent bonds thus create free electrons and holes. Hence the intrinsic carrier density increases significantly and becomes higher than the impurity density. In this case, Si no longer behaves as semiconductor hence this temperature is the upper limit of its operation temperature.

As discussed at the beginning of this chapter, carrier mobility is affected by the scattering effects. Different scattering effects have different temperature dependencies. When temperature rises, the scattering caused by the acoustic phonons decreases while the scattering caused by the ionised impurities increases [24]. Furthermore, the increasing intrinsic carrier density due to higher temperature will lead to a higher chance for scattering between carriers. The accurate description for carrier mobility is complicated hence the basic empirical equations are used, as shown in Eqns. 2.69 and 2.70 [44]. The decrease of carrier mobility

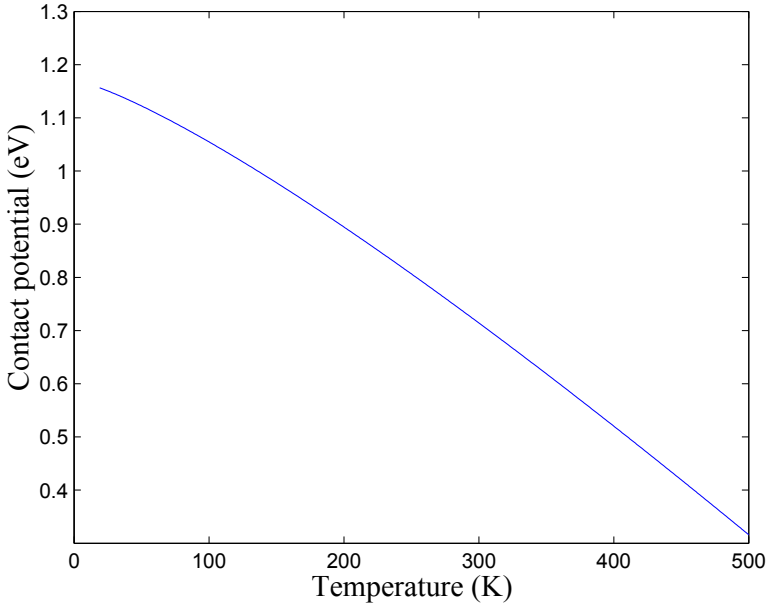


Figure 2.16: Contact potential changes with temperature.

will lead to the drop of conductivity according to Eqn. 2.21. In the case of high-level injection, when the intrinsic carrier density is negligible, the increase of temperature will lead to a higher on-state voltage drop. The electron mobility also affects the transconductance coefficient of MOS channel K_{PL}, which can be calculated from Eqn. 2.71. Here, w and l are the channel width and length respectively. In the case of device modelling, simplified Eqn. 2.72 is used to describe the K_{PL} relationship with temperature [20]. The decrease of K_{PL} will lead to the decrease of MOS channel current when V_{GE} and V_{TH} is kept constant according to Eqn. 2.59.

$$\mu_n = 1400 \left(\frac{300}{T}\right)^{2.5} \tag{2.69}$$

$$\mu_p = 450 \left(\frac{300}{T}\right)^{2.5} \tag{2.70}$$

$$K_{PL} = \frac{w\mu_n C_{OX}}{l} \tag{2.71}$$

$$K_{PL} = K_{PL}(300) \cdot \left(\frac{300}{T}\right)^{0.8} \tag{2.72}$$

Auger recombination is the dominant mechanism that limits the carrier lifetime under high-level injection. The high-level injection carrier lifetime due to this mechanism is given by Eqn. 2.11. According to the results given in [45], R_{bc} decreases with the increasing temperature. Therefore, τ_{HL} rises when temperature increases or free carrier density decreases. τ_{HL} is a very important parameter that affects both the device on-state and switching characteristics. Higher τ_{HL} means more charge are stored in the drift region during on-state. This could lead to a lower on-state resistance. However, the trade-off is a higher reverse recovery current which will lead to the increase of IGBT switching time and limits the maximum allowed switching frequency. A simplified equation is used to describe the temperature dependency of τ_{HL} according to [20], as shown in Eqn. 2.73.

$$\tau_{HL} = 5 \times 10^{-7} \left(\frac{T}{300}\right)^{1.5} \tag{2.73}$$

The threshold voltage can be calculated according to [46], as shown in Eqn. 2.74. Here, Q_D is the charge stored per unit area in the inversion layer, Q_{SS} is the surface charge trapped in the gate oxide, and $\Delta\psi_{Th}$ is the potential difference across the depletion layer in the bulk semiconductor at threshold voltage or above. Q_D and $\Delta\psi_{Th}$ are given by Eqns. 2.75 and 2.76, respectively. Taking all temperature dependent variables into account and applying the common design parameters: $N_a = 1 \times 10^{17}\ cm^{-3}$, $N_d = 1 \times 10^{20}\ cm^{-3}$, $C_{OX} = 3.45 \times 10^{-8}\ F/cm^2$, $Q_{SS} = 4 \times 10^{-8}\ C/cm^2$, the relationship between threshold voltage and temperature can be plotted as shown in Fig. 2.17. It is clear that V_{TH} decreases with the increasing temperature. In the case of device modelling, a simplified relationship is used by linear approximation, as shown in Eqn. 2.77 [46]. Here $V_{TH}(300)$ is the threshold voltage

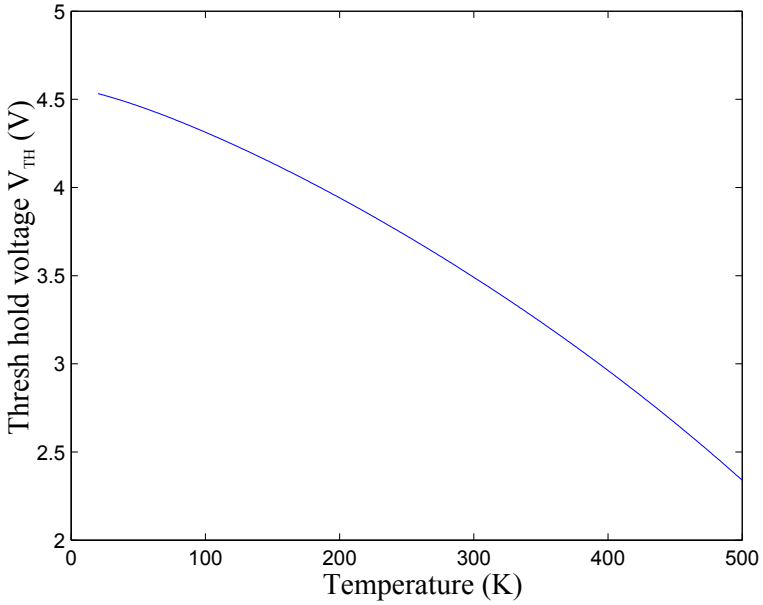


Figure 2.17: Threshold voltage changes with temperature.

at 300 K. The decrease of V_{TH} could lead to a shorter turn-on delay time and consequently longer turn-off delay time. It also affects I_{MOS} according to Eqn. 2.59, when IGBT is operating in saturation region. The effect of temperature on I_{MOS} is then complicated since the opposite temperature dependency of V_{TH} and K_{PL} hence needs to be considered according to different conditions.

$$V_{TH} = \frac{Q_D - Q_{SS}}{C_{OX}} + \Delta\psi_{Th} \tag{2.74}$$

$$Q_D = \sqrt{2q\varepsilon N_a \Delta\psi_{Th}} \tag{2.75}$$

$$\Delta\psi_{Th} = \frac{2kT}{q} ln(\frac{N_a}{n_i}) \tag{2.76}$$

$$V_{TH} = V_{TH}(300) - 6 \times 10^{-3}(T - 300) \tag{2.77}$$

Chapter 3

Physics of Power Module Package Failures

3.1 Introduction to Packaging Technology for Power Modules

Before the 1970's, the semiconductor devices in a power system had to be mounted to separate heatsinks and isolated from each other, which made the power electronic systems large and expensive. The invention of power modules in the mid of 1970's made it possible to combine different devices in one plastic housing and share one heatsink. This is done by soldering their common terminals together with electrical contacts on a direct copper bonded (DCB) substrate. The DCB substrate generally consists of three layers: two metal layers (copper) with a thin ceramic layer (e.g. Al_2O_3) between them. It provides good thermal conductivity and electrical isolation between the chips and heatsink. With such power module design, the power electronic systems can be made more compact, cost efficient, and reliable.

The package of a power module refers to the substrate which the chip is mounted on, the electrical connections between chips within the module, and the housing. The package not only provides terminals for the chips to connect with the external circuit, it also protects

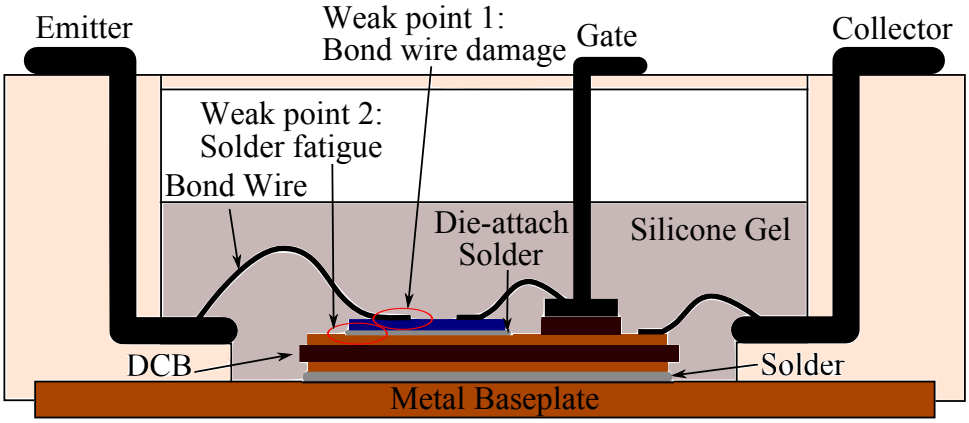


Figure 3.1: Cross section view of a conventional power module.

Table 3.1: Materials and their CTE used for each layer of the conventional power module.

Layer	Material	CTE (ppm/°C)
Bond wire	Aluminum	22
Die	Silicon	3
Die-attach solder	SnAg(3)	28
DCB-copper	Copper	17.5
DCB-ceramic	Al_2O_3 or AlN	7 or 4
DCB-copper	Copper	17.5
DCB-solder	SnAg(3)	28
Baseplate	Copper or AlSiC	17.5 or 8

them from the external environment and provides mechanical support. An ideal package should be able to provide these electrical connections without any parasitics, extract heat efficiently and be robust. However in reality, there is always a compromise.

Fig. 3.1 shows the cross section view of a typical conventional u-package power module proposed by Mitsubishi [47]. The IGBT chip is soldered onto a DCB and connected to the terminals via bond wires. The DCB is mounted onto a thick metal baseplate to enhance the mechanical strength of the module. The electrical circuit in the power module is covered by silicone gel to protect it from air or other corrosives. This structure consists of many layers of different materials; the materials for each layer and the associated coefficient of thermal expansion (CTE) are listed in Table 3.1.

Parasitic problems due to the package have been widely investigated [48, 49]. The most severe parasitic is the inductance caused by the internal copper leads and bond wires. IGBT turn-off voltage overshoot is proportional to this circuit stray inductance (L_s). Furthermore, IGBT turn-on and turn-off speed are also reported to be affected by L_s [50]. Therefore, it is desirable to find a better solution to lower the module parasitic inductance. Many recent investigations have reported reduced module internal inductances, using ribbon bonds [51], metal post interconnections [52], and solder bump interconnections [53].

Power semiconductor chips produce massive amounts of heat during switching and when conducting. Therefore, efficient cooling techniques are required for power modules. It can be seen from Table 3.1 that the CTE mismatch between the DCB ceramic layer and the copper baseplate is large. This will cause fatigue of DCB solder when the module is thermal cycled. AlSiC baseplate [54] can be used to significantly reduce this fatigue problem since the CTE of AlSiC matches that of the DCB as shown in Table 3.1. However, its low thermal conductivity and high price make it undesirable for the design of cost efficient power modules. When the power module is mounted to the heatsink, thermal grease must be applied either to the heatsink or to the module baseplate to displace any air gaps between them. The thermal conductivity of such grease is quite low, hence must be as thin as possible. The practical thickness for the grease between baseplate and heatsink is approximately 100 μm. This reduces the efficiency of the heat extraction significantly. Considering these disadvantages of the baseplate, the power module without baseplate is presented by Semikron [54]. This design not only eliminates the DCB solder fatigue problem but also reduces the thickness of the thermal grease to 20 μm since the DCB is more flexible than metal baseplate. Therefore, better thermal conductivity is achieved. Another advantage of this design is the reduction of the the IGBT module weight and cost.

It is clear that heat can only be extracted from the bottom of the chip for the conventional power module. Stockmeier proposed a double sided cooling structure to enhance the cooling capability of the power module [55], as shown in Fig. 3.2. The chip is soldered between two

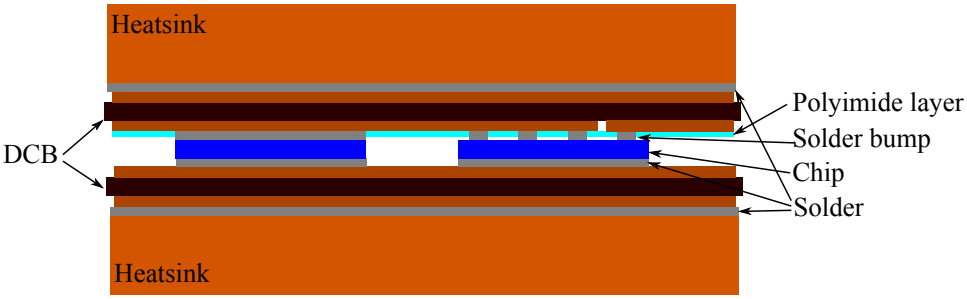


Figure 3.2: Cross section view of the double sided cooling IGBT module.

DCBs which are directly brazed on heatsinks, hence the thermal resistance is minimised. Flip-chip solder bumps [56] are proposed to replace the wire bonds on the top of the IGBT chip. The top DCB is chemically etched to ensure the insulation between gate and emitter connections for the IGBT chip. A microchannel heatsink is also proposed in that paper to improve the efficiency of the heat extraction.

The reliability challenges of power module are mainly based on the thermomechanical stress caused by the CTE mismatch between adjacent layers when the module suffers temperature cycles. In the case of conventional power modules, the weak points appear at the two interfaces shown in Fig. 3.1, since they suffer large CTE mismatch and the largest temperature cycles. Consequently, the most commonly observed failure modes for IGBT module are bond wire lift off and solder fatigue [17, 57]. Researchers have made great efforts to improve the reliability of packaging. The use of ribbon bonds as shown in Fig. 3.3 not only reduces the parasitic inductance but also enhances the bond wire reliability [51]. One ribbon bond can substitute several bond wires therefore the time needed for the bonding process is less and the manufacturing cost can be reduced. However, the bonding of ribbon requires higher pressure on the chip, which means that thicker chip metallisation is necessary [51]. The double sided cooling power module eliminates the bond wire, but the reliability of the flip-chip solder bumps needs to be studied. Sinter technology has been proposed for the die bond to overcome the solder fatigue problem under high temperature cycles [58]. The reliability for traditional soft solder (lead-tin)is poor when the operation temperature is above

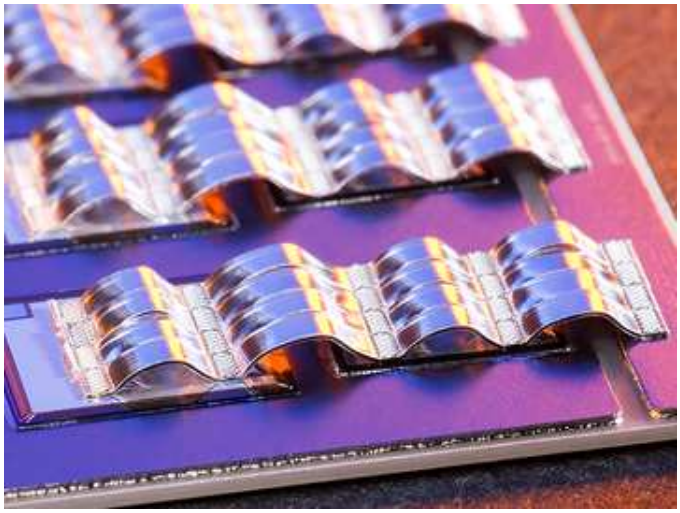


Figure 3.3: Ribbon bonds [1].

Table 3.2: Materials and their CTE used for each layer of the conventional power module.

Properties	Solder layer - SnAg(3)	Silver sinter layer
Melting point (℃)	221	962
Thermal conductivity (W/K)	0.07	0.24
Tensile strength (MPa)	30	55

125 ℃ due to their relatively low melting temperature. Silver is a good replacement material for solder to mount the die onto the DCB. The comparison of these two materials is shown in Table 3.2. With the silver sinter technique, the maximum operation temperature (close to 200 ℃) for Si can be achieved. It is a desired technique for power module that incorporate SiC semiconductor devices since the operation temperature of SiC can be much higher than Si. Also, the higher tensile strength of Silver ensures better power cycle ability for silver sinter. High pressure is applied to lower the sinter temperature in order to prevent the chip from high sinter temperature (400 ℃) [59].

3.2 Reliability Estimation via Physics of Failure Approach

Fatigue damage accumulation is an old but unsolved problem. The fatigue damage was first seriously considered in the mid-nineteenth century when a widespread failure of railway axles were observed [60]. More than 50 fatigue damage models have been proposed during the last century to predict the lifetime of materials under cyclic loading [61].

As discussed in chapter 1, in order to obtain an accurate lifetime prediction for IGBT modules, it is essential to understand the physics of different failure mechanisms. However, none of the currently available physics-based models are able to obtain an accurate prediction for all applications due to the complexity of the fatigue processes. The applicability of each model varies under different conditions. Therefore, empirical models extracted from experimental results are still widely used for reliability design purposes. The detailed discussion of the fatigue damage process is beyond the scope of this thesis. In this section, the basic physical processes of the two major failure mechanisms are discussed.

3.2.1 Basic Physics of Bond Wire Lift Off

Wire bonding is the process of providing electrical connection between the silicon chip and the external leads of the power module. The most widely used wire bonding technique is the aluminum (Al) wedge wire bonding which uses ultrasonic energy and a downward pressure for a specific duration to form the wedge bond between the Al wire and a metalized bond pad. During this process, the Al wire is plastically deformed due to the applied bonding force and ultrasonic energy. This leads to a high dislocation density and strain energy at the bonding interface. Dislocation is the carrier of deformation, just like the electron is the carrier of charge [3]. After the applied force and ultrasonic energy is removed, the temperature of the bonding interface will drop to room temperature. A residual stress is then produced due to the thermal contraction and CTE mismatch between Al and Si. The maximum stress

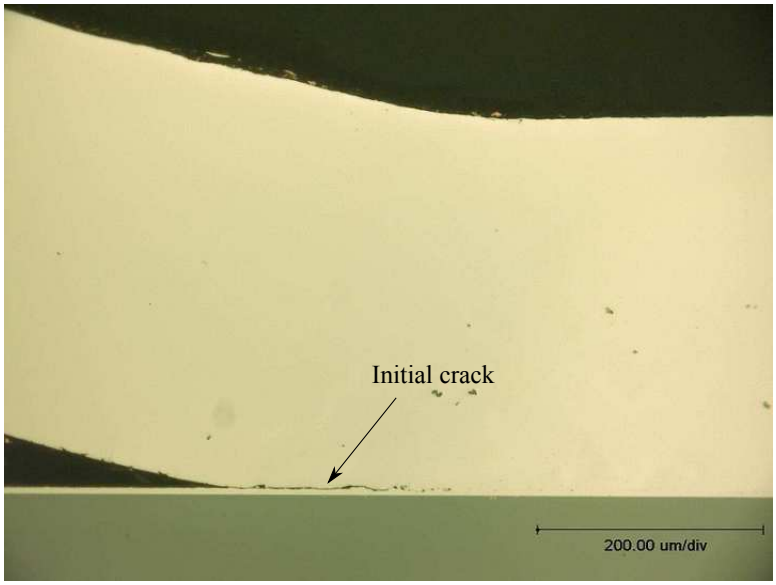


Figure 3.4: Initial crack at the heel of bond wire [2].

appears at the heal and toe of the bond wire, hence the crack initiates from the ends of the bond foot and travels to the center along the plane above the bonding interface [2]. Consequently, the initial cracks are often found at the bond heel due to the residual stress and flexing of wire during the bonding process, as shown in Fig. 3.4. Bond wire lift off occurs when the cracks from two ends meet at the center.

Fig. 3.5 illustrates the crack propagation by ductile tearing in the bond wire. When a shear stress is applied between the bond wire and die, the stress close to the crack (σ_{local}) is greater than the average stress (σ) applied to the bond wire due to the stress concentrating effect. The local stress can be calculated according to Eqn. 3.1 [3], where a is the crack length and r is the distance to the crack tip as defined in Fig. 3.5. It is clear that the local stress decreases when its distance to the crack tip increases. The plastic flow occurs when the local stress is greater than the yield stress of Al. The plastic zone width (r_y) can be calculated by setting $\sigma_{local} = \sigma_y$ in Eqn. 3.1 and is shown in Eqn. 3.2. Within the plastic zone, plastic flow takes place around the deformation caused by wedge bonding process or the impurities in Al wire and forms individual voids. As the plastic flow propagates, these

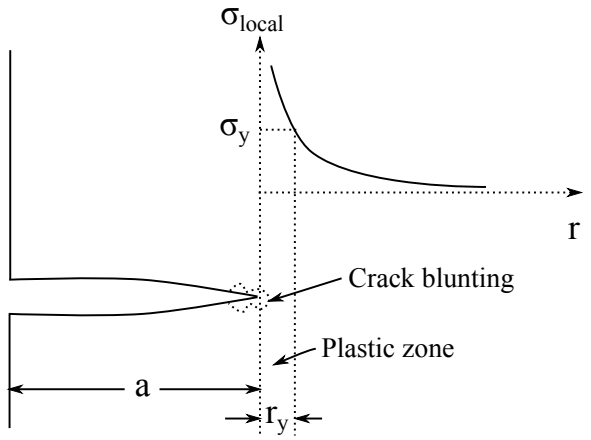


Figure 3.5: Crack propagation by ductile tearing [3].

voids link up and turn the initial sharp crack into a blunt crack, as shown in Fig. 3.5. This kind of plastic flow is termed ductile tearing [3], which leads to the decrease of σ_{local}. Furthermore, ductile tearing consumes a lot of energy by plastic flow and causes the rough surface of the crack, as shown in Fig. 3.6.

$$\sigma_{local} = \sigma + \sigma\sqrt{\frac{a}{2r}} \tag{3.1}$$

$$r_y = \frac{a\sigma^2}{2(\sigma_y - \sigma)^2} \tag{3.2}$$

There have been several attempts to estimate the lifetime of bond wires by numerical simulation based on its physics. Ramminger presented FEM models including grain boundaries in 1998 [62], while Hager presented the lifetime estimation method based on computational plasticity in 2000 [63]. However, both models are limited by the uncertainty in evaluating the initial stresses induced by the plastic deformation during the wedge bonding process. Therefore, the commonly accepted estimation of bond wire reliability is still based on experiments such as accelerated tests.

The Coffin-Manson law is widely used to estimate the material lifetime under plastic strain, as shown in Eqn. 3.3, where N_f is the number of cycle to failure under the plastic

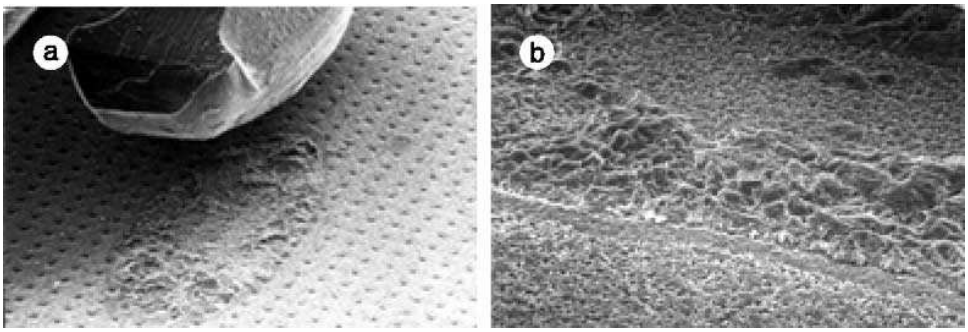


Figure 3.6: (a) Scanning electron microscope (SEM) image of bond wire lift off, (b) footprint of Al bond wire after lift off [4].

strain ϵ_{pl}, a and b are constants fitted from experimental data. The plastic strain caused by the thermomechanical stress at the interface of the Al wire and Si chip can be calculated from a simple two layer structure according to Eqn. 3.4, where α_{Al} and α_{Si} are the CTE of Al and Si respectively, ϵ_{el} and ϵ_{pl} are the elastic and plastic strain respectively, L and h are the length and thickness of the bond foot respectively. When ΔT_j is large, the total strain (ϵ) can be considered as ϵ_{pl}, hence the Coffin-Manson law can be rewritten in the form of Eqn. 3.5.

$$N_f = a\epsilon_{pl}^{-b} \tag{3.3}$$

$$\epsilon = \frac{L}{h}(\alpha_{Al} - \alpha_{Si})\Delta T_j = \epsilon_{el} + \epsilon_{pl} \approx \epsilon_{pl} \tag{3.4}$$

$$N_f = a\left(\frac{L}{h}(\alpha_{Al} - \alpha_{Si})\Delta T_j\right)^{-b} = a'\Delta T_j^{-b} \tag{3.5}$$

$$\epsilon = \frac{\sigma}{G} \tag{3.6}$$

The following example demonstrates the validity of the Coffin-Manson plastic deformation assumption. According to the text book [3], the Shear modulus of Al (G_{Al}) is 26 GPa,

the yield strength of Al (σ_{yAl}) is 40 MPa, the yield shear strain can be estimated by Hooke's Law as shown in Eqn. 3.6, and the result is 1.54×10^{-3}. The length/thickness ratio of the bond foot is around 5 for the power module (SKM50GB123D) under investigation in this work. The CTE mismatch between Al and Si is 19 ppm/K according to Table 3.1. By substituting these values into Eqn. 3.4, the temperature swing (16 ℃) needed for plastic deformation can be calculated. Under field operating conditions, a temperature cycle of 16 ℃ can be easily achieved. In the case of accelerated testing, the temperature cycle is even greater, for example 80 ℃ or 120 ℃ [57]. Therefore the plastic deformation assumption holds and the Coffin-Manson Law can be used for the lifetime estimation of bond wire lift off.

3.2.2 Basic Physics of Solder Fatigue

Solder is widely used in the power semiconductor industry to form the electrical and mechanical connection between the chip and substrate. Therefore, the ideal solder should have excellent electrical properties and mechanical strength. Also, the melting point of solder should not too high to damage the chip during soldering process. The most frequently used solders in power modules are based on tin-silver, indium, or tin-lead alloys [4]; these are soft materials with low melting temperature and high electrical conductivity. The chip consumes energy and generates heat when it is switching and conducting current. A potentially large amount of heat needs to conduct from the chip to the heatsink, through the solder and DCB, therefore a high thermal conductivity is desirable.

During the operation of power module, temperature cycles are generated due to the change of load conditions. Cyclic thermomechanical stress is induced at the solder layer due to the temperature cycle and CTE mismatch between the Si chip and DCB substrate. This cyclic stress will cause solder fatigue and possibly failure. Fatigue is defined as the process of progressive failure caused by repeated stress cycles which are well below the yield strength of the material [3]. The propagation of a fatigue crack is similar to the crack propagation process and is shown in Fig. 3.7. The thermomechanical shear stress produces a plastic zone

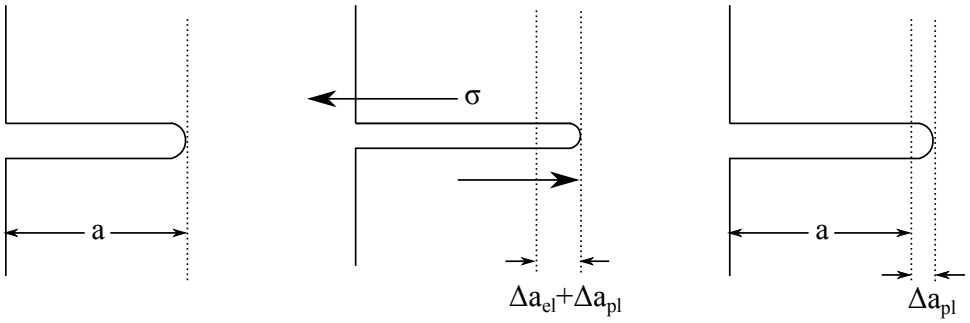


Figure 3.7: Illustration of fatigue crack propagation.

at the fatigue crack tip and causes the crack elongate by Δa and create a new surface. Δa consists of the elastic deformation (Δa_{el}) and plastic deformation (Δa_{pl}). Δa_{el} is released after the stress is removed. However, the new surface extends the original crack by Δa_{pl} due to the permanent plastic deformation. Fig. 3.8 shows the hysteresis loop of this process for one stress cycle.

There are many models available to describe the solder fatigue damage [64]. They can be organised into three categories: strain based models [65], energy based models [66, 67], and damage based models [68]. Energy based models estimate the lifetime by calculating the overall stress-strain hysteresis energy of the solder joint. The damage based models do the same job by calculating the accumulated damage caused by crack propagation. These two types of model require a FEA tool to obtain the stress-strain hysteresis loop hence are time consuming. The strain based models predict failure from calculated or experimentally determined shear strain. The Coffin-Manson model is one of the best known strain based models, as shown in Eqn. 3.3. Thermally induced strain at the solder layer can be derived from a three layer structure [69], as shown in Fig. 3.9.

The maximum stress appears at the edge and can be calculated using Eqn. 3.7, with β as defined by Eqn. 3.8. Here, G stands for the shear modulus and E stands for the Young's modulus of the related layer. The die and substrate are much more rigid than the solder, hence $\beta L_{die} \ll 1$ and $tanh(\beta L_{die}) \approx \beta L_{die}$. Therefore, the maximum shear strain at the edge

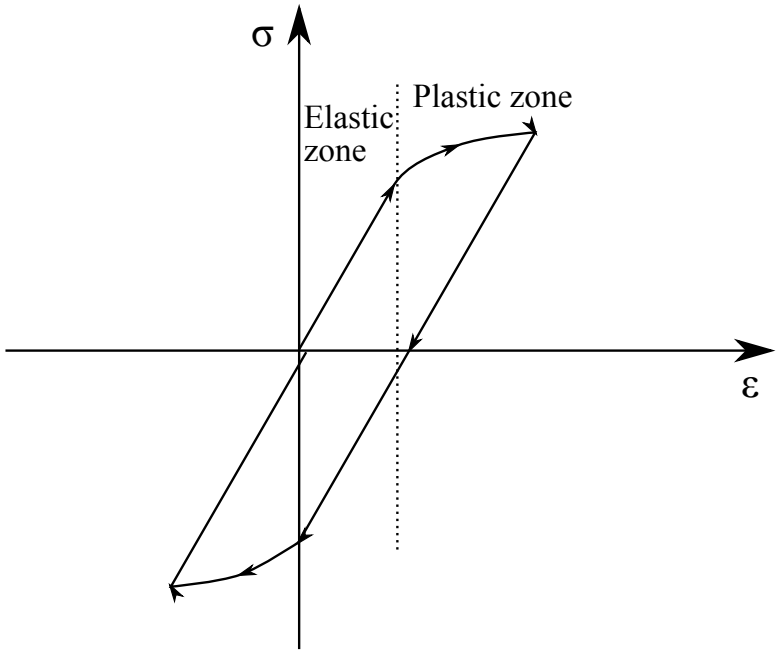


Figure 3.8: Hysteresis loop for the solder fatigue process for one stress cycle.

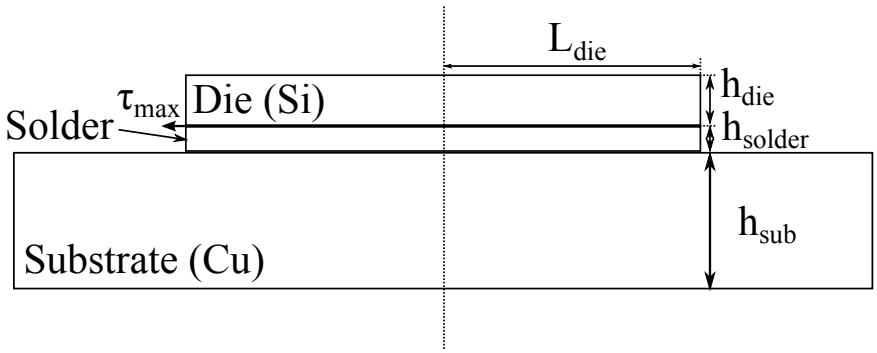


Figure 3.9: Three layer structure of the electronic device assembly.

of the solder joint can be expressed in the form of Eqn. 3.9. The lifetime is then linked with temperature cycles by substituting Eqn. 3.9 into the Coffin-Manson Law, as shown in Eqn. 3.10.

$$\sigma_{max} = \frac{G_{solder}}{h_{solder} \cdot \beta} \cdot (\alpha_{sub} - \alpha_{die}) \cdot \Delta T_j \cdot tanh(\beta L_{die}) \tag{3.7}$$

$$\beta^2 = \frac{G_{solder}}{h_{solder}} \cdot \left(\frac{1}{E_{sub} h_{sub}} + \frac{1}{E_{die} h_{die}} \right) \tag{3.8}$$

$$\epsilon_{max} = \frac{L_{die}}{h_{solder}} \cdot (\alpha_{sub} - \alpha_{die}) \cdot \Delta T_j \tag{3.9}$$

$$N_f = a \left(\frac{L_{die}}{h_{solder}} \cdot (\alpha_{sub} - \alpha_{die}) \cdot \Delta T_j \right)^{-b} \tag{3.10}$$

This is the same as the bond wire failure, it needs to be examined whether the assumption of Coffin-Manson Law holds. According to Ochiai [70], the shear modulus of solder changes with temperature. It decreases from 10 Gpa at room temperature (25 ℃) to 5 GPa at 100 ℃. The minimum value is used to obtain the largest possible yield strain. The yield strength of SnAg solder is around 30 MPa according to [71]. The yield strain of this solder can be calculated using Eqn. 3.6, giving a value of 6×10^{-3}. The typical thickness of the die-attach solder layer is from 50 to 100 μm; the average value is used in this calculation. The length of the chip of the selected module is 4 mm and the CTE mismatch between Si and Cu is 14.5 ppm/K. Substituting these values into Eqn. 3.9 gives the minimum ΔT_j needed for plastic deformation as 8 ℃. This is even lower than that required for the plastic deformation of Al wire, thus the assumption of Coffin-Manson Law holds.

The solder fatigue process is complex. Some papers state that the lifetime is not only depends on the temperature cycle amplitude (ΔT_j), but also the mean temperature (T_m), the frequency of the cycle (f), and the dwelling time (t_d) [72, 73]. Norris and Landzberg

modified the conventional Coffin-Manson relationship and proposed a model which considers f, maximum junction temperature (T_{jmax}) and ΔT_j as variables [74], as shown in Eqn. 3.11. Here, E_a is the activation energy and k is the Boltzmann's constant. However, more experimental data is needed in order to extract the parameters of this model, which leads to a longer investigation and higher cost.

$$N_f = A \cdot f^B \cdot \Delta T_j^a \cdot exp\left(\frac{E_a}{kT_{jmax}}\right) \tag{3.11}$$

3.3 Brief Introduction to Other Power Module Failures

Other failure mechanisms of power modules have also been observed by researchers, although not as frequent as the two main failures [4, 75]. They are briefly reviewed in this section.

Bond wire heel cracking is mainly observed after a long endurance test especially when the wedge bonding process is not optimised [4]. This failure is also caused by the thermo-mechanical stress during temperature cycles. When the temperature increases, the wire will expand and flex, generating a stress at the bond heel which leads to the fatigue. However, it is reported that this failure mechanism is slower than the bond wire lift off even when the bonding process is not optimised.

Die fracture failure may occur during temperature cycles when defects within the Si chip are present. Microcracks might be developed in the manufacturing processes such as crystal growth, wafer scrubbing and slicing [75]. Tensile and shear stresses are generated when the chip exercises temperature cycles. Hence the manufacturing defect may develop into a crack and grow gradually by fatigue propagation, which may eventually lead to a die fracture failure. Brittle fracture of the die can occur suddenly without any plastic deformation when the initial defect exceeds a critical size.

The IGBT gate oxide also faces reliability problems. Gate oxide slow trapping could occur after long term operation due to the diffusion of electrons into the silicon-oxide interface. The

presence of electrons shifts the gate threshold voltage and decreases the switching speed of the device, which may eventually cause system failures in high frequency operation applications. This type of failure also needs to be considered in high temperature applications since the electron trapping process could be accelerated by increasing the temperature or electric field [9]. The gate oxide might breakdown at low electric fields after a long period, which is called the time dependent dielectric breakdown (TDDB) [76]. Some research shows that long term operation at high temperatures (e.g. 250 ℃) is not practical for gate oxide [77] while others report the opposite [78]. Therefore, more investigation is needed in this area.

Corrosion is more frequently observed in early generation of power modules where the Al wires and chip metallisation are exposed to air. A thin oxide layer grows on the surface of the metals. Current power modules contain silicone gel to protect the metal from the environment. However modules that operate in harsh environments still suffer damage from corrosion, therefore special design is needed [9].

Chapter

4 Inverter Electrothermal Model

As discussed in chapter 1, the junction temperature (T_j) is one of the key parameters that affect the IGBT module lifetime. However, T_j is difficult to measure directly during the inverter operation since it is buried in the module package. This chapter introduces an inverter electrothermal model implemented in PLECS which is a tool for high speed simulations of power electronic systems in Matlab/Simulink. The aim of this model is to simulate T_j during operation with a certain mission profile. To speed up the simulation the inverter model interpolates power dissipation from LUTs rather than running the device models each time switching losses are required, which are relatively slow. The LUTs are generated by running the IGBT and diode models [20, 79] under different operating conditions. The thermal networks are extracted from the heating waveforms of IGBT junction and case temperatures which are measured on a power cycling rig. After putting together all this information, T_j can then be calculated with the power dissipation and thermal networks according to the heat conduction equation.

4.1 Parameterisation of Device Switching Model

An important reason for circuit simulation in power electronic systems is to estimate the device power losses under different operating conditions. The instantaneous switching power loss of a device is far higher than its conduction power loss. Therefore the switching loss cannot be ignored under high frequency operation although the switching period is short. Consequently, the power loss cannot be simply calculated by $P = V \cdot I$. Calculating switching losses require an accurate physics based device model, since lots of device parameters are highly temperature dependent as discussed in chapter 2. Furthermore, a careful parameter extraction sequence is needed to ensure the simulation results match the experimental curves at different operating conditions.

A two-step parameter extraction method for the device models is described in [41] and its accuracy at different temperatures is shown in [22]. In the first step of the extraction procedure, an inductive switching test is required to record the IGBT and diode switching waveforms at different temperatures. Six diode parameters and twelve IGBT parameters (for NPT IGBTs) are required in order to create an accurate simulation. In the second step, an optimisation procedure is designed to refine the parameters. A summary of the extraction process is shown below:

1. Measure the device switching waveforms from inductive switching test.
2. Initial estimation based on the device datasheet, empirical value range and the measured switching waveforms.
3. Simulation of the device switching model using the estimated parameters.
4. Compare the simulated waveforms with the measured waveforms and calculate the error between them.
5. Modify the parameters to minimize the error.

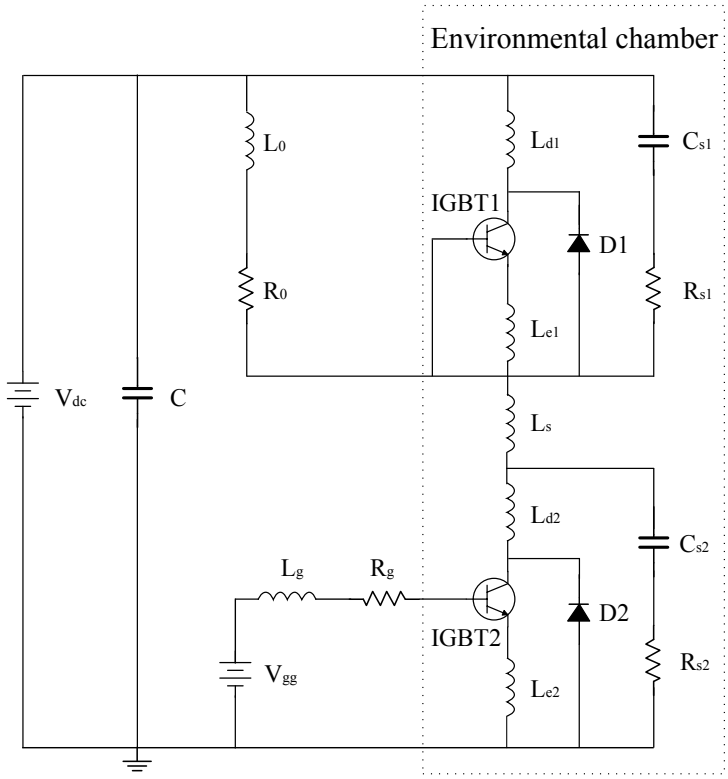


Figure 4.1: Circuit diagram for the inductive switching test.

4.1.1 Inductive Switching Test

Some parameters, such as carrier lifetime, need to be extracted from the device switching waveform. Therefore, an inductive switching test is designed to record the switching waveforms of both IGBT and diode at different temperatures. Fig. 4.1 shows the detailed circuit diagram for the inductive switching test. A half-bridge IGBT module SKM50GB123D is selected for the investigation. One IGBT module is placed in an environmental chamber where temperature can be controlled from -75 ℃ to 200 ℃ (Tenney TUJR environmental test chamber). When IGBT2 is switched on, current flows through the inductor L_0, stray inductance L_s, diode series inductance L_d, IGBT2 and Kelvin emitter inductance L_e. When IGBT2 is switched off, the inductor L_0 is discharged via Diode1.

A digital phosphor oscilloscope (Tektronix TD55054B) is used to record the switching

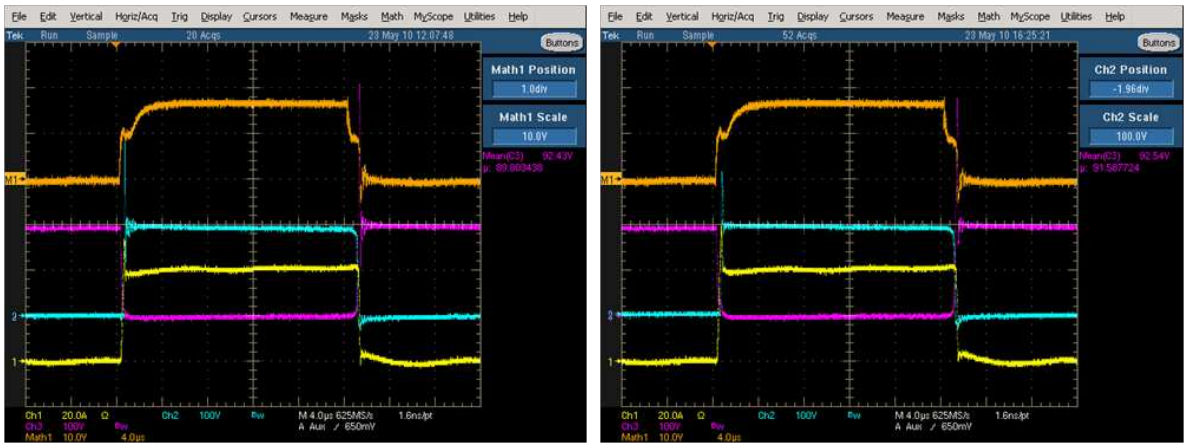


Figure 4.2: Inductive switching waveforms at 30 ℃ (left) and 130 ℃ (right).

waveforms of the IGBT collector-emitter voltage drop (V_{CE}, purple), IGBT gate-emitter voltage (V_{GE}, orange), voltage across diode1 (V_{AK}, cyan) and collector current (I_C, yellow).

The operation parameters of the inductive switching test are listed below:

$R_g = 27\ \Omega$, $I_C = 40\ A$, $V_{GG} = 17\ V$, $L_0 = 1.78\ mH$, $V_{DC} = 200\ V$.

The temperature of the environmental chamber is changed from 30 ℃ to 130 ℃ and the switching waveforms are recorded every 20 ℃. Fig. 4.2 shows the examples of recorded switching waveforms at different temperatures. The waveforms recorded by the digital oscilloscope are saved in data files for result analysis. These waveform data are then imported and filtered in Matlab. More comparative figures are plotted to focus on the temperature effects on the switching waveforms, as shown in Fig. 4.3. The effect of increasing temperature can be summarised below:

- Both diode turn-off voltage overshoot and its slope dV_{AK}/dt are reduced.
- Both IGBT turn-off voltage overshoot and its slope dV_{CE}/dt are reduced.
- IGBT turn-on current overshoot is increased while its slope dI_C/dt is not affected.
- IGBT turn-off delay time is longer.

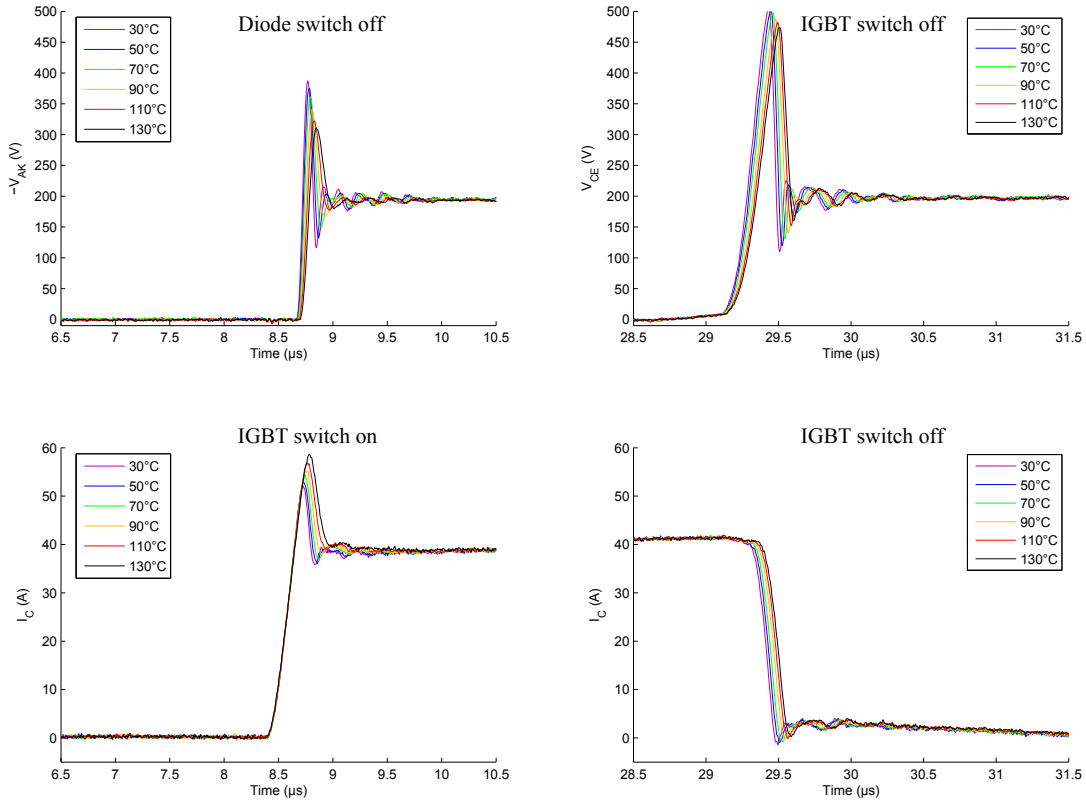


Figure 4.3: Comparison of switching waveforms at different temperatures.

When the temperature increases, the Si intrinsic carrier density increases exponentially according to Eqn. 2.2. The effective doping concentration increases with temperature as well [80]. Both of effects lead to a higher carrier density in the N- drift region. The rate of carrier extraction is constant since it is determined by the circuit inductance. Therefore a longer time is needed to extract the stored carriers which leads to a greater IGBT turn-on current overshoot. The depletion layer in diode N- drift region builds up slower since it takes more time to push the stored charge out. Consequently, both the dV_{AK}/dt and V_{AK} overshoot are smaller.

In the case of IGBT turn-off, more free carriers are stored in the N- drift region. Similar to diode turn-off, it takes more time to push the stored charge out. The depletion layer in the IGBT N- drift region builds up slower hence both dV_{CE}/dt and V_{CE} overshoot are smaller. Furthermore, the threshold voltage V_{TH} decreases according to Fig. 2.17. Therefore, it takes more time for V_{GE} to drop below V_{TH}, which leads to a longer turn-off delay time.

4.1.2 Initial Parameter Extraction of Device Switching Model

The goal of the parameter extraction is to fit the simulation results with the experimental waveforms, especially V_{AK}, V_{CE} and I_C, since they are used to calculate the power loss. There are three parameter extraction methodologies employed in the initial extraction procedure: estimation based on empirical value range, extrapolation according to device datasheet and extraction based on device switching waveforms. Table 4.1, 4.2 and 4.3 summarise the initial parameter extraction methods for diode, IGBT and circuit respectively.

Table 4.1: Initial parameter extraction of diode.

Symbol (unit)	Description	Extraction method
$A\ (cm^2)$	Active die area	$A = I_F/J$, where I_F is the average forward DC current and J is the maximum current density of diode, normally between 100 and 150 $A{\cdot}cm^{-2}$
$\tau_{HL}\ (\mu s)$	High-level lifetime	$\tau_{HL} = Q_{RR}/I_F$, where Q_{RR} is the reverse recovery charge
$W_d\ (\mu m)$	Drift region width	$V_{DB} = bW_d/ln(aW_d)$, where V_{DB} is the break down voltage, a and b are constants, $a = 1.07 \times 10^6\ cm^{-1}$ and $b = 1.65 \times 10^6\ V \cdot cm^{-1}$
$N_B\ (cm^{-3})$	Drift region doping	Normally between 6×10^{13} and $2\times10^{14}\ cm^{-3}$
$h_n\ (cm^4 \cdot s^{-1})$	Electron recombination coefficient	Initial estimation of $1 \times 10^{-14}\ cm^4 \cdot s^{-1}$
$h_p\ (cm^4 \cdot s^{-1})$	Hole recombination coefficient	Initial estimation of $1 \times 10^{-14}\ cm^4 \cdot s^{-1}$

Table 4.2: Initial parameter extraction of IGBT.

Symbol (unit)	Description	Extraction method
$V_{TH}\ (V)$	MOS threshold voltage	Extract from forward IV characteristics or use the value given in datasheet
$K_{PL}\ (A \cdot V^{-2})$	MOS transconductance coefficient	Extract from forward IV characteristics as shown in Fig. 4.4: $K_{PL} = I_{MOS}/V_{DS}^2 = b \cdot I_c/[(b+1)V_{DS}^2]$, where b is the ratio of electron and hole mobilities
$\lambda\ (V^{-1})$	Short channel parameter	Extract from forward IV characteristics as illustrated in Fig. 4.4
$C_{GE}\ (nF \cdot cm^{-2})$	Gate-emitter capacitance	$C_{GE} \approx C_{ies}$, where C_{ies} is the IGBT input capacitance

Continued on next page

Table 4.2 – continued from previous page

Symbol (unit)	Description	Extraction method
C_{OX} (nF)	Oxide capacitance	$C_{OX} = max(C_{res})/(A \cdot a_i)$, where C_{res} is the IGBT reverse capacitance
A (cm^2)	Effective die area	Measured by opening device package or estimated from the formula: $A = I_{CM}/J$, where I_{CM} is the peak collector current from the RBSOA curve given in the datasheet and J is the maximum current density of IGBT, normally between 100 and 250 $A \cdot cm^{-2}$
a_i	Ratio of inter-cell to total die area	$a_i \approx min(C_{res})/min(C_{oes})$, where C_{oes} is the IGBT output capacitance
N_B (cm^{-3})	Doping concentration of N- drift region	Typical N_B value for simulation is 1×10^{14} cm^{-3} or $N_B = 2.88 \times 10^{17} \times V_{BR}^{-1}$, where V_{BR} is the breakdown voltage from datasheet plus 150 to 200 V typical margin for general IGBTs
W_d (μm)	Drift region width	$W_d = (\epsilon \cdot E_c)/(q \cdot N_B)$ (for NPT IGBT) or $W_d = \epsilon(E_c - \sqrt{E_c^2 - 2q \cdot N_B \cdot V_{BR}/\epsilon})/(q \cdot N_B)$ (for PT IGBT), where E_c is the Critical electrical field value for silicon, Normally between 2×10^5 and 3×10^5 $V \cdot cm^{-1}$, q is the unit electron charge $(1.6 \times 10^{-19}\ C)$, ϵ is the permittivity of silicon $(1.04 \times 10^{12}\ F \cdot cm^{-1})$

Continued on next page

Table 4.2 – continued from previous page

Symbol (unit)	Description	Extraction method
τ_{HL} (μs)	Carrier high-level life time in N- drift region	Extracted from the turn-off current tail, it is estimated as the time constant of exponential current decay curve. Fig. 4.5 shows the life-time extraction region of the turn off current tail.
τ_{BF} (μs)	Carrier lifetime in N buffer layer (PT)	Similar to high-level life time extraction. The effective lifetime of PT IGBTs change with clamp voltage, therefore it need to perform under several clamp voltages. Fig. 4.6 shows the lifetime extraction for PT IGBTs.
h_p $(cm^4 \cdot s^{-1})$	Electron recombination coefficient in emitter (NPT)	Initial value is set to be 1×10^{-12} $cm^4 \cdot s^{-1}$. Increase h_p reduces the level of stored charge and therefore reduce $I_c(0+)$ in Fig. 4.5.
W_H (μm)	Width of N buffer layer (PT)	Typical value is 4 to 10 μm
N_H (cm^{-3})	Doping concentration of N buffer layer (PT)	Typical value is 10^{16} to $10^{17} cm^{-3}$
I_{sne} (A)	Minority carrier saturation current (PT)	Typical value is 10^{-14} to 10^{-12} A

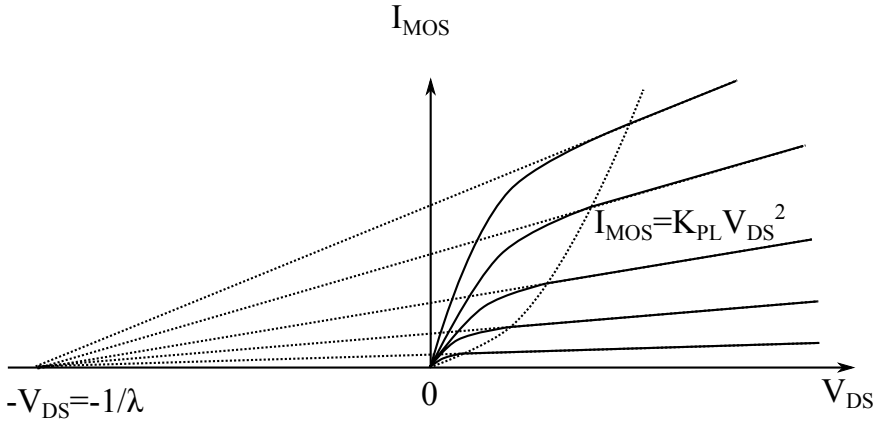


Figure 4.4: Forward V-I characteristics of IGBT for parameter extraction.

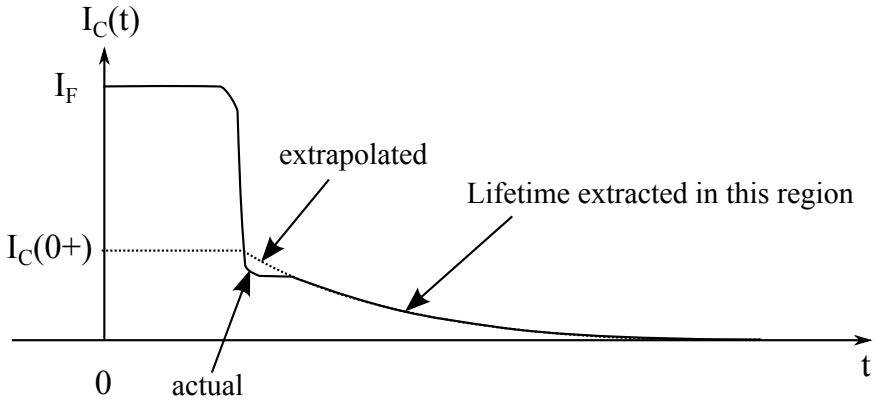


Figure 4.5: IGBT turn off current waveform.

4.1.3 Optimisation of the Extracted Parameters

An optimisation procedure is needed to refine the initial extracted parameters and obtain a better fit between simulation and experimental waveforms. It contains three steps: run the circuit simulation with estimated parameters, compare the simulated waveforms with the experimental waveforms and calculate the error figure, modify the estimated parameters to minimize the error.

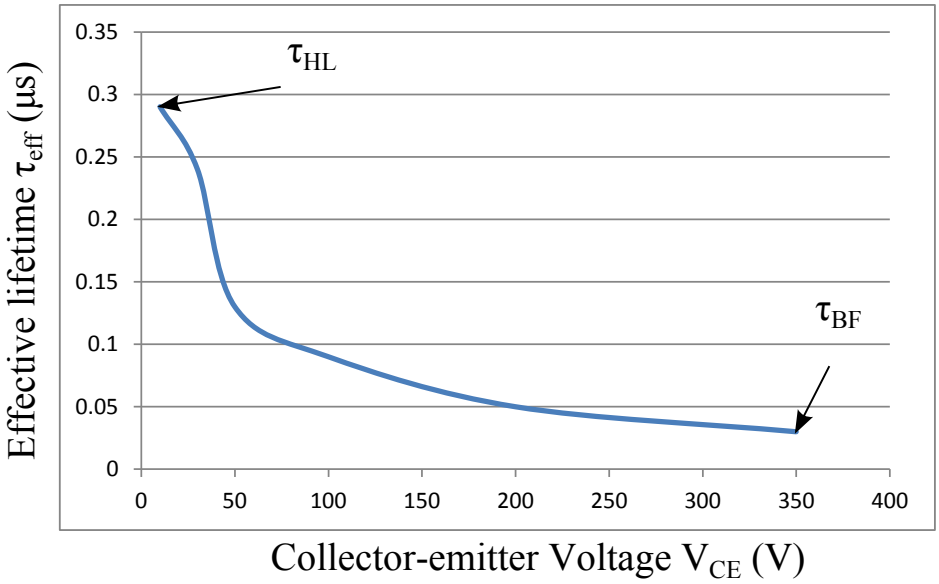


Figure 4.6: Effective lifetime extraction under different clamped voltage for PT IGBT.

Table 4.3: Initial parameter extraction of diode.

Symbol (unit)	Description	Extraction method
L_s (H)	Circuit stray inductance	$L_s = (V_{CM} - V_{DC})dt/dI_c$, where V_{CM} is the peak IGBT collector voltage during turn-off, V_{DC} is the DC supply voltage
L_e (H)	Kelvin emitter inductance	Typical value is 5 to 20 nF for 1 cm^2 IGBT
τ_{load} (s)	Load time constant	$\tau_{load} \approx R_0 \times L_0$

The goal of the optimisation process is to estimate the switching loss accurately. A possible method is to compare the salient points, such as IGBT turn on voltage overshoot, diode reverse recovery time. With only small amount of points to match, the computation time of this method is shortened by sacrificing the accuracy. A much more accurate method is to calculate the sum of squared errors at each time point between simulation and experimental waveforms. Normalized errors are used to avoid the unit mismatch between different parameters, such as current and voltage. The synchronisation of the waveforms is required to avoid the errors in di/dt and dv/dt. It is carried out by aligning the gate drive signals,

which are square waves, of the simulation and experiment.

The error figure may be based on either instantaneous power dissipation waveforms or current and voltage waveforms. One disadvantage of the former method is the loss of independence of the current and voltage waveforms, since the same power can be maintained by current increasing and voltage decreasing and vise versa. Matching voltage and current waveforms will tend to ensure the instantaneous power dissipation waveform is matched as well. Therefore the sum of the normalized squared errors for the current and voltage waveforms at both turn-on and turn-off is selected as the error figure, as shown in Eqns. 4.1 and 4.2.

$$f_e = SSE(\frac{V_{CE,on}}{V_{DC}}) + SSE(\frac{I_{C,on}}{I_F}) + SSE(\frac{V_{CE,off}}{V_{DC}}) + SSE(\frac{I_{C,off}}{I_F}) \tag{4.1}$$

$$SSE(x) = \sum_{n=1}^{N}(x_{sim}(n) - x_{meas}(n))^2 \tag{4.2}$$

Finally, the optimisation process can be summarised below:

1. Run the simulation with the initial extracted parameters and calculate the base error f_e.

2. Run the simulation at each surrounding point ($\pm\Delta x_i$ for each parameter x_i) and calculate the error.

3. Find the lowest error, if it is lower than the base error, set it as the new base point; else the minimum error appears at the current base point and the search is terminated.

4. Change the same parameter in the same direction as in step 3, and evaluate the error. If the error is still lower, this becomes the new base point; else return to step 2.

Fig. 4.7 shows the simulation results (solid line) compare with experiment waveforms (dashed line) at both 30 ℃ and 130 ℃. It is obvious that the simulation waveforms fit the

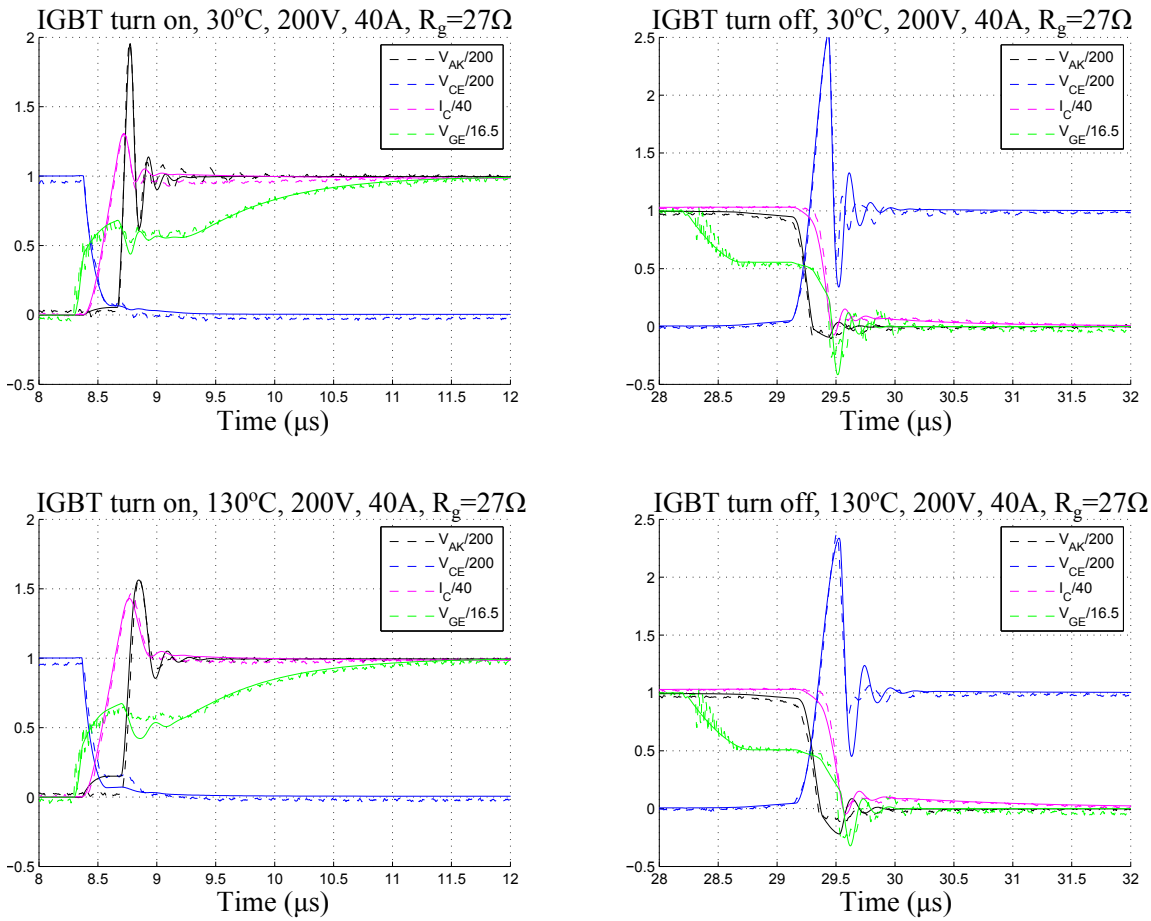


Figure 4.7: Comparison of simulation and experimental switching waveforms at different temperatures.

experimental results well. The main contribution of errors are the turn off voltage oscillations and V_{CE} turn on plateau, which may be due to the simplification of the device model (only the drift region part is physically based).

4.1.4 Look Up Tables for IGBT and Diode Power Loss

LUTs are used in large time scale inverter simulations to speed up the computation. With accurate device switching models, the IGBT and diode switching power loss LUTs can be built by running the simulation at different voltages, currents and temperatures. However, the simulation only gives the instantaneous power dissipation. The switching loss is the integration of the instantaneous power dissipation over the switching region. Therefore it is essential to find out the switching start and end points. The following algorithm is used, its Matlab code is available in Appendix C.2.

a. Find IGBT turn-on/off start points t_{i1}/t_{i3}, and on/off state stable point mid1/mid2 which are fixed time stamp defined in the simulation.

b. Find the maximum extreme point max_ext between t_{i1} and mid1, and compare each maxima point ext(x) between max_ext and mid1 with on-state power loss $sP_i(mid1)$, if $ext(x) \leq 1.5 \times sP_i(mid1)$ then IGBT turn-on end point t_{i2} is ext(x); else, it is the first point $sP_i(t_{i2})$, which meets the requirement, between the last maxima point and mid1.

c. Use the same method to find IGBT turn-off end point and diode turn-off/on end point.

Fig. 4.8 shows an example of locating the IGBT turn-off end point. The red cross indicates the maxima point of the power curve, the green cycle indicates the located IGBT turn off end point by the above algorithm. The remaining switching end points of IGBT and diode can be located by following the same algorithm. Run the device switching model at different voltages (100 V, 200 V, 300 V, 400 V), currents (10 A, 20 A, 30 A, 40 A, 50 A)

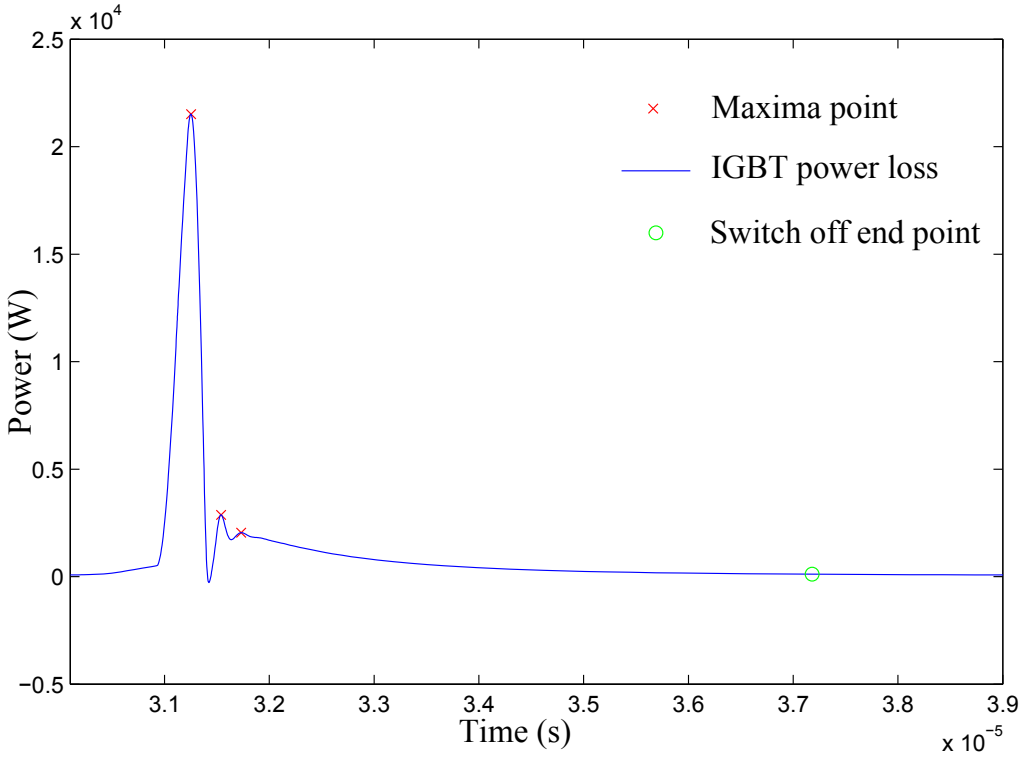


Figure 4.8: Demonstration of the algorithm for determine the switching region.

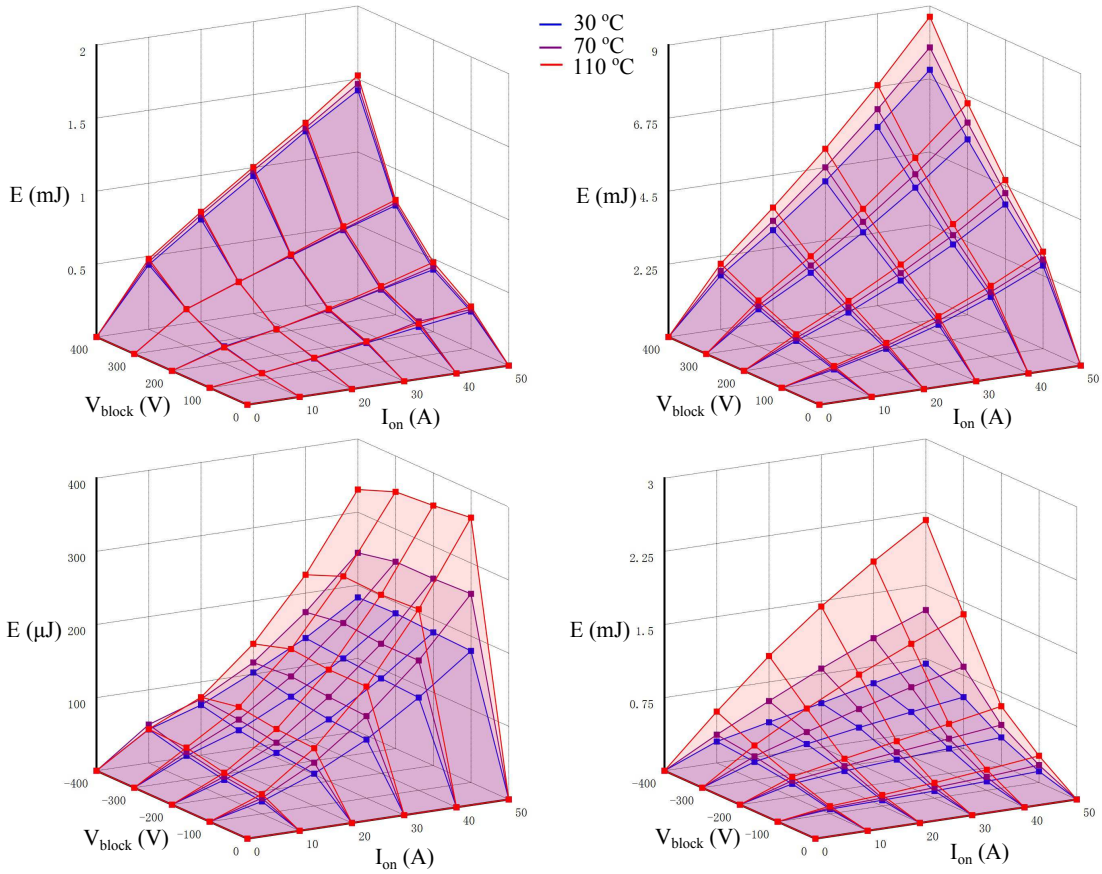


Figure 4.9: IGBT and diode switching loss LUTs. Top-left: IGBT turn-on loss; top-right: IGBT turn-off loss; bottom left: diode turn-on loss; bottom right: diode turn-off loss.

and temperatures (30 ℃, 70 ℃, 110 ℃). The switching loss (E_{on}/E_{off}) can be calculated by integrating the power loss between t_{i1}/t_{i3} and t_{i2}/t_{i4}. Fig. 4.9 shows the LUTs of IGBT and diode switching loss.

IGBT and diode on-state loss can be calculated from their forward IV characteristics. Three IGBT modules are randomly selected and placed in the environmental chamber. The IV characteristics of these IGBTs and diodes are measured by a curve tracer at different temperatures (30 ℃, 70 ℃, 110 ℃). Fig. 4.10 shows the IV characteristics for both IGBT and diode at different temperatures.

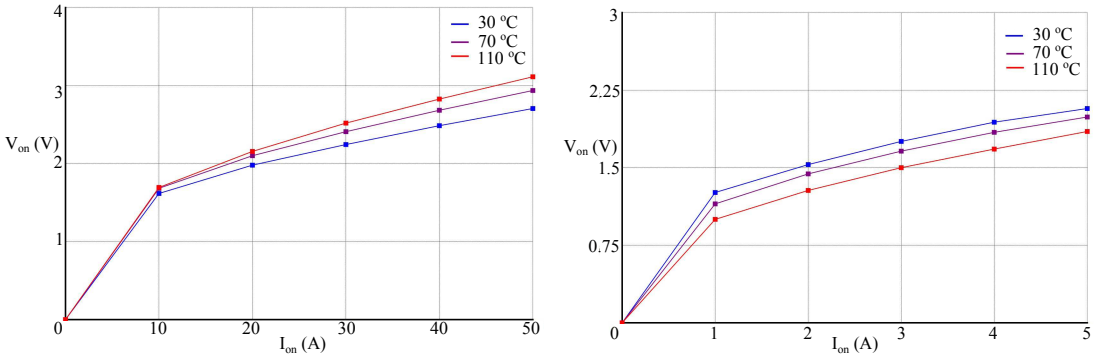


Figure 4.10: IGBT (left) and diode (right) on-state power loss LUTs

4.2 Thermal Network Extraction for IGBT Module and Heatsink

As discussed in chapter 2, temperature affects the behavior of power semiconductor devices greatly, since many of their properties are strongly temperature dependent. Therefore thermal models for an IGBT module and heatsink are required in order to simulate the device junction temperature during inverter operation. In this section, the principles of heat propagation are introduced, followed by a practical thermal network extraction and optimisation method.

4.2.1 Principles of Heat Propagation

The propagation of heat in a system can take place in three different ways, convection, radiation or conduction. In solid materials, such as electronic devices, conduction is the dominant heat transfer method. Eqn. 4.3 describes the heat conduction in a homogeneous isotropic material [81]. The heatflow is assumed to be one-dimensional for simplification.

$$\frac{\partial^2 T}{\partial x^2} = \frac{c \cdot \rho}{\lambda_{th}} \cdot \frac{\partial T}{\partial t} \tag{4.3}$$

Where, λ_{th} is the heat conductance, c is the thermal capacitance, ρ is the density of the

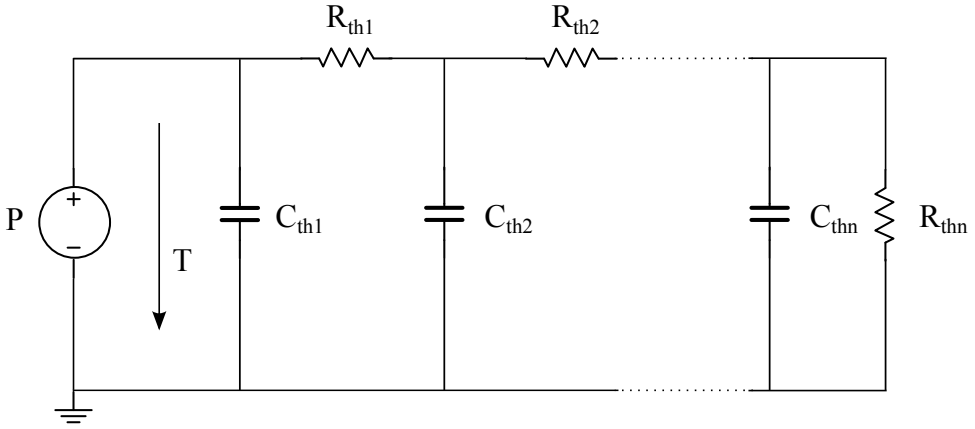


Figure 4.11: Electrical transmission line equivalent circuit diagram (Cauer network) for modeling heat conduction properties.

material, T is the temperature and x is the coordinate in the direction of heat propagation.

In the case of heat conduction in solid media, there is no direct comparison for the electrical term of inductance, and a volume element cannot cool itself. This can be described formally by both inductance and transverse conductance equal to zero. Applying these restrictions to the transmission line equation [81], it can be simplified in the form of Eqn. 4.4, which has the same structure as heat conduction equation.

$$\frac{\partial^2 V}{\partial x^2} = C \cdot R \cdot \frac{\partial V}{\partial t} \tag{4.4}$$

Where, V is the voltage, x is the coordinate in the direction of current flow, C is the capacitance per unit length and R is the resistance per unit length.

According to Kirchhoff's research [81], "Two different forms of energy behave identically when the basic differential equations which describe them have the same form and the initial and boundary conditions are identical". Therefore, the heat conduction process can be modeled by a transmission line style equivalent circuit diagram which consists of RC elements only, as shown in Fig. 4.11. The equivalences between electrical and thermal variables are listed in Table 4.4.

Table 4.4: Comparison of equivalences between thermal and electrical variables.

Thermal		**Electrical**	
Temperature	T in K	Voltage	U in V
Heat flow	P in W	Current	I in A
Thermal resistance	R_{th} in $K \cdot W^{-1}$	Resistance	R in Ω
Thermal capacitance	C_{th} in $J \cdot K^{-1}$	Capacitance	C in F

The Cauer network, although it is mathematically complicated, refers to the physics of equivalent circuit diagram of heat conduction and is the only network that is able to describe the internal temperature distribution of the system correctly. In this work, only the IGBT and diode junction temperatures are of interest, hence another equivalent circuit diagram can be used, such as the Foster network shown in Fig. 4.12. Each individual RC element of Foster network represents one term of a partial fractional division of the thermal transfer function of the system, whereby the order of the individual terms is arbitrary [81]. It is clear that Foster network is mathematically simple and the thermal impedance can be described by Eqn. 4.5:

$$Z_{th}(t) = \sum_{i=1}^{n} R_{thi}(1 - e^{-\frac{t}{R_{thi} \cdot C_{thi}}}) \tag{4.5}$$

A two RC element Foster network is selected in this work for simplicity. All thermal power generated by the IGBT chip is assumed to conduct from junction to case. Therefore the thermal impedance Z_{th} can be extracted from the heating curves of junction T_j and case T_c temperature by Eqns. 4.6 and 4.7:

$$\frac{T_j(t) - T_c(t)}{P(t)} = \sum R_{thi}(1 - e^{\frac{-t}{R_{thi} \cdot C_{thi}}}) \tag{4.6}$$

$$P(t) = V_{CE}(t) \cdot I_c \cdot D \tag{4.7}$$

where, $V_{CE}(t)$, $T_j(t)$ and $T_c(t)$ are measured during a designed test, D is the duty ratio of the power supply and I_C is the IGBT collector current.

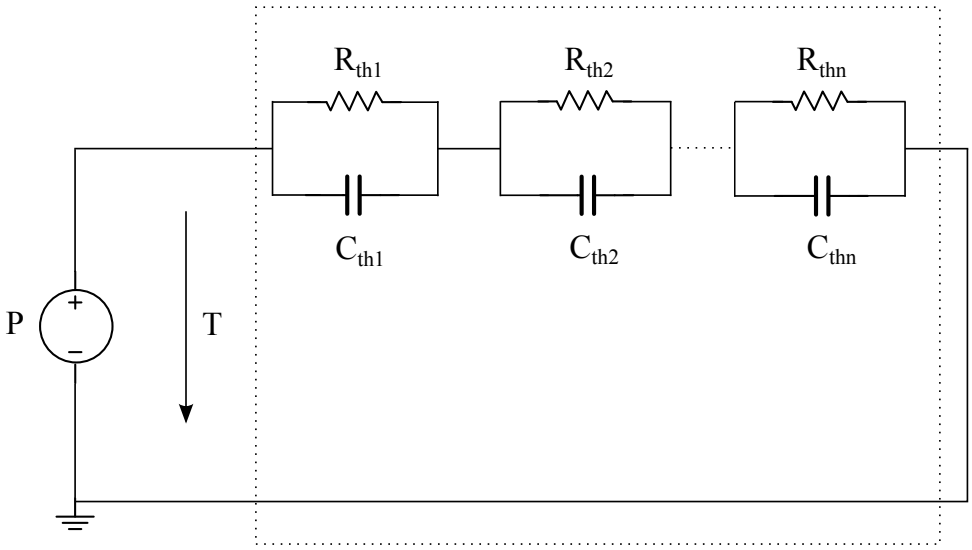
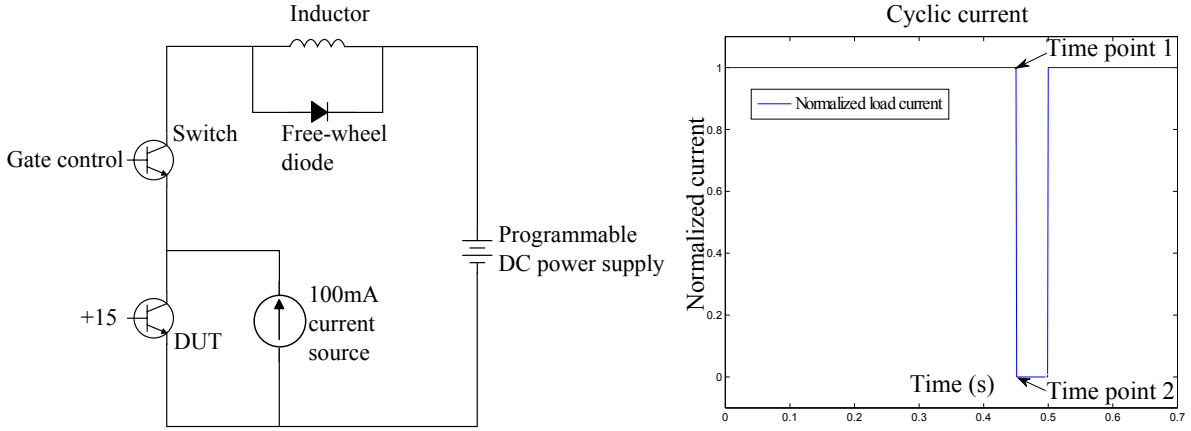


Figure 4.12: Electrical transmission line equivalent circuit diagram (Foster network) for modeling heat conduction properties.

4.2.2 Initial Thermal Network Extraction

A test is designed to measure the temperature and voltage curves for the thermal impedance extraction, its circuit diagram is shown in Fig. 4.13(a). IGBT forward voltage drop V_{CE} at 100 mA is selected as a temperature sensitive parameter to sense IGBT junction temperature since it decreases linearly with increasing temperature [82]. A cyclic load current is supplied to the DUT by turning off the switch every 0.5 s since T_j can only be measured when $I_C = 100$ mA. T_j, T_c, and water temperature at the inlet (T_{win}) and outlet (T_{wout}) of the heatsink are recorded at time point 2 of each load cycle, as shown in Fig. 4.13(b). This process takes about 0.05 s, therefore the duty ratio D is 0.9. V_{CE} at load current is measured and recorded at time point 1 of each load cycle, as shown in Fig. 4.13(b). The heating phase terminates when the system is thermally stable. The detailed control flow chart diagram for the thermal impedance extraction test is shown in Appendix D.1. Fig. 4.14 shows the temperature and voltage curves of one IGBT recorded in this test, as well as its fitted thermal impedance.

The thermal network can also be extracted from cooling curves. The switch is turned off



(a) Thermal network extraction test circuit (b) Cyclic load current

Figure 4.13: Thermal network extraction

during the cooling phase, hence T_j can be recorded continuously in this phase. However in this case, as the heat conducts from junction to case, P_{jc} cannot be calculated electrically. Therefore the heat carried away by water P_{water} is used to estimate P_{jc} indirectly. The thermal impedance can be extracted from Eqns. 4.8 and 4.9. The recorded temperature curves and extracted thermal impedance are shown in Fig. 4.15.

$$\frac{T_j(t) - T_c(t)}{P(t)} = \sum R_{thi} \cdot e^{\frac{-t}{R_{thi} \cdot C_{thi}}} \tag{4.8}$$

$$P_{jc} \approx P_{water} = C_{water} \cdot m_{water} \cdot \Delta T = C_{water} \cdot \rho_{water} \cdot V_{water} \cdot (T_{wout} - T_{win}) \tag{4.9}$$

Comparing the results fitted from the above two extraction methods, the former appears to be more accurate. The thermal networks extracted from both methods are based on the temperature differences between junction and case. The temperature difference during heating (30 ℃ as shown in Fig. 4.14) is much greater compared with that during cooling (2 ℃ as shown in Fig. 4.15). Furthermore, the power estimated by the second method is based on the temperature difference between the input and output water, which is also relatively small. Ideally, both methods should come to the same result. However, noise is unavoidable

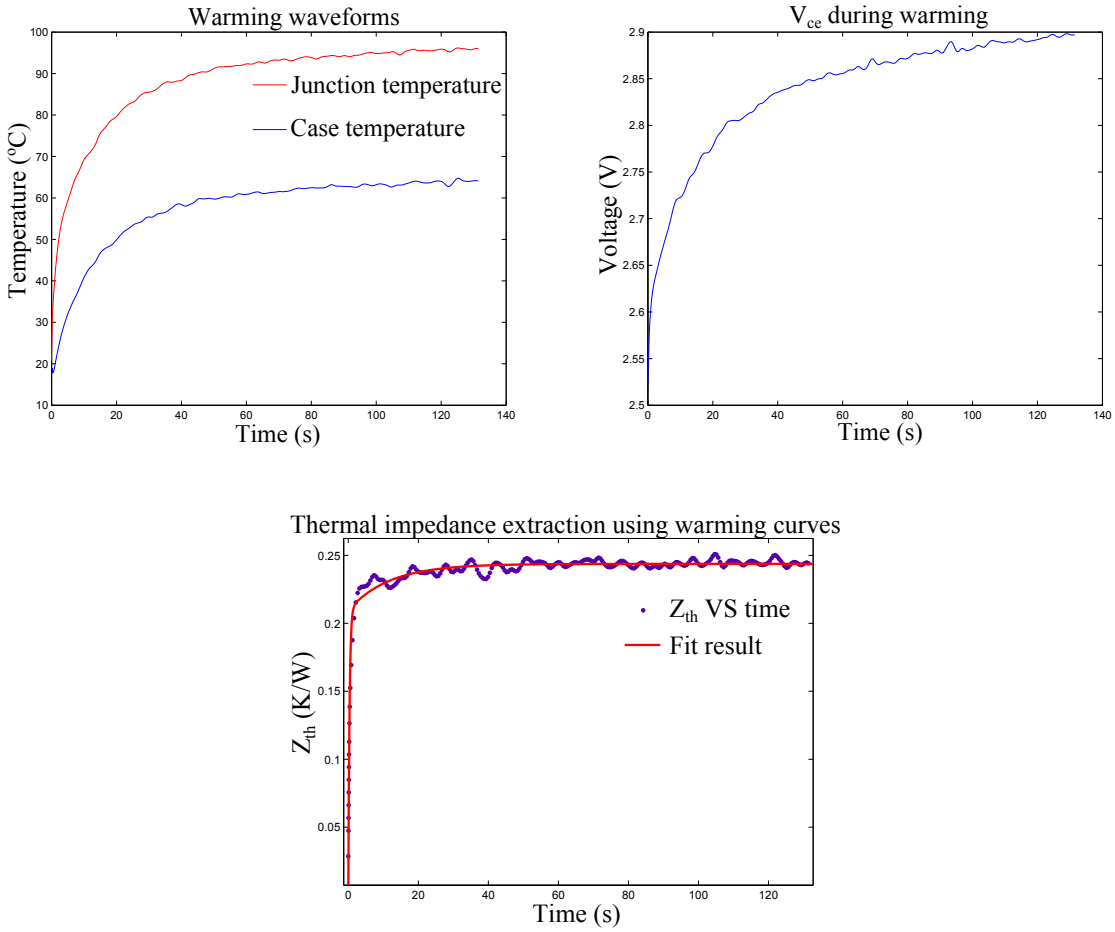


Figure 4.14: Junction to case thermal impedance extraction: heating curves (up two) and fitting result (down).

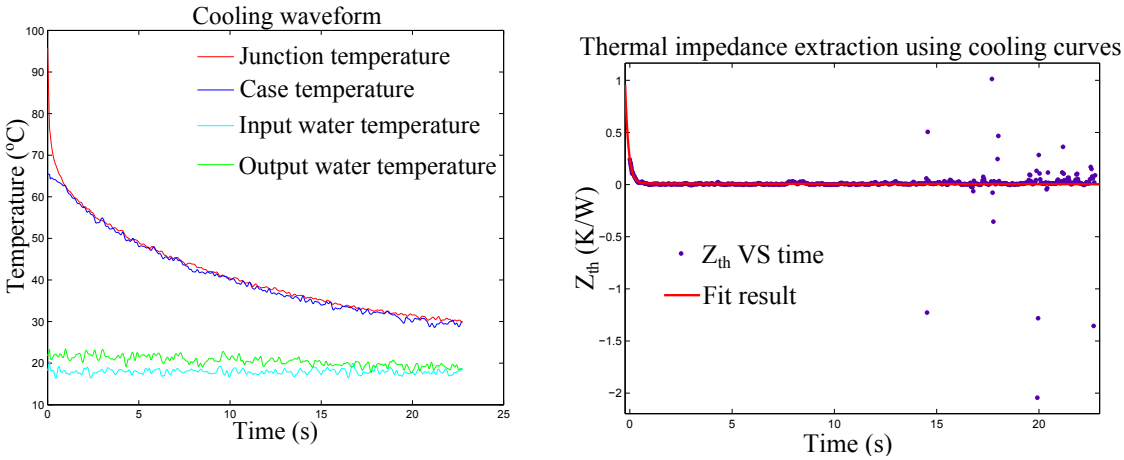


Figure 4.15: Junction to case thermal impedance extraction: cooling curves (left) and fitting result (right).

for experimental measurements. In the case of thermal couples, the measurement noise is roughly ±2 ℃. Therefore the affect of noise on the thermal network extracted from cooling curves is much greater than that on the heating curves.

Fig. 4.16 shows the R_{th} of 22 IGBT modules extracted from both the heating curve and the cooling curve. The results illustrate that the R_{th} extracted from heating curves are greater than that from cooling curves. This is because the heat carried away by water during cooling is the sum of power discharged from the thermal capacitance between junction and case and the power discharged by the heatsink. Hence $P_{water} > P_{jc}$ which consequently leads to a smaller R_{th}.

4.2.3 Optimisation of Extracted Thermal Network

Both methods introduced in section 4.2.2 introduce errors from assumptions, measurements, filtering and curve fitting. A careful optimisation process is needed to approach a more accurate simulation. An electrothermal model is built in PLECS to simulate the thermal network extraction test, as shown in Fig. 4.17. The IGBT is fed with a constant 45 A current during heating phase to keep the power dissipation the same as that in the experiment. The

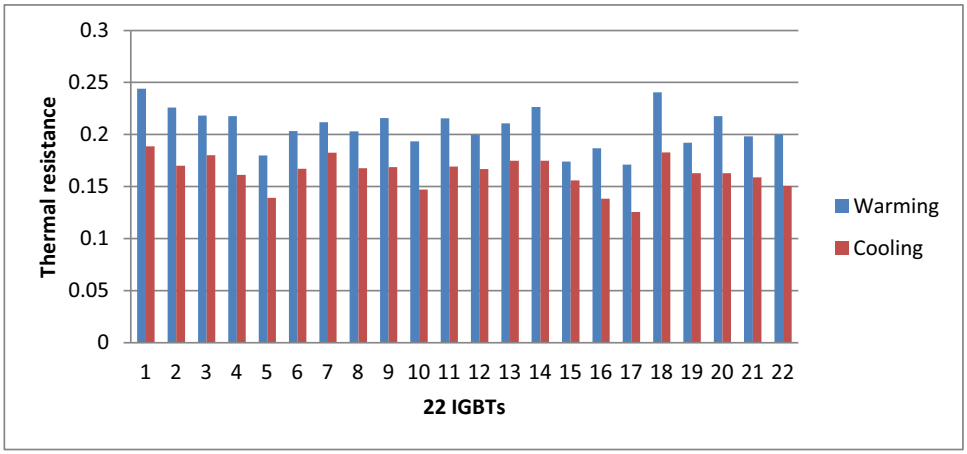


Figure 4.16: Comparison of two thermal impedance extraction methods.

IGBT and diode power loss LUTs and initial extracted thermal networks are used to describe the electrical and thermal behavior of the module. The Thermal Chain block, which is a build in block in PLECS, consists of a two-level Cauer network and is used to represent the thermal behavior of the heatsink. The extracted network from experiment is Foster, hence a transformation is necessary. This can be done with the build in function in PLECS. T_{varG} is set to be the average water temperature. The optimisation process is similar to that described in section 4.1.2. Each element of the thermal network of the IGBT module and heatsink is modified to locate the minimum error figure between simulation results and experimental waveforms. Fig. 4.18 shows the simulation results before and after the optimisation. Clearly, better fitting between simulation and experimental waveforms is achieved after optimisation. The extracted Cauer network before and after optimisation for both IGBT module and heatsink is listed in Table 4.5. Here, t_1 and t_2 indicates the time constant for each level of RC network. It need to be noted that the change of the time constant may due to the measurement error or due to the inadequate of thermal network levels. A more complected thermal model can be used to increase the accuracy of the simulation.

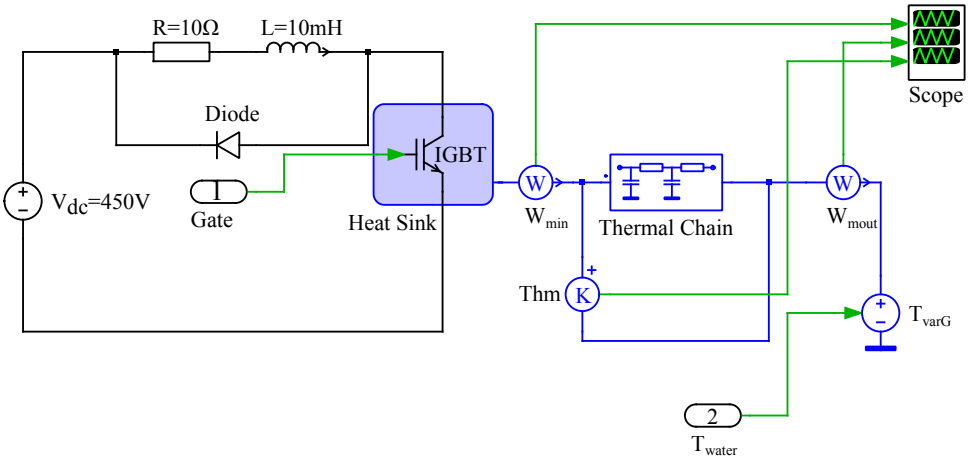


Figure 4.17: Simulation of thermal impedance extraction test

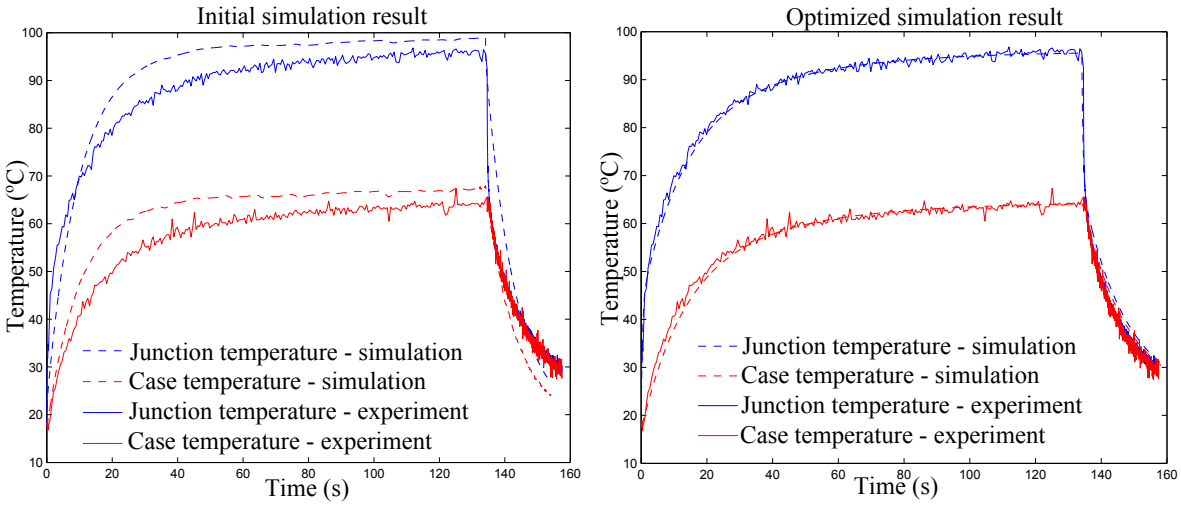


Figure 4.18: Simulation of thermal impedance extraction test

Table 4.5: Thermal networks before and after optimisation.

Cauer network		$\mathbf{R_1}$ **(K/W)**	$\mathbf{C_1}$ **(J/K)**	$\mathbf{t_1}$ **(s)**	$\mathbf{R_2}$ **(K/W)**	$\mathbf{C_2}$ **(J/K)**	$\mathbf{t_2}$ **(s)**
IGBT	Initial	0.04733	2.136	0.101	0.1965	12.88	2.492
	Optimized	0.22	1.8	0.396	0.02	36	0.72
Heatsink	Initial	0.2964	0.2063	0.061	0.05873	3.195	0.188
	Optimized	0.32	0.2	0.064	0.02	800	16

4.3 Inverter Electrothermal Model in PLECS

There are many different topologies currently available for inverter design. Multilevel power inverters have been receiving increased attention in the past few years for high power applications [83, 84]. The main advantage of this type of topology is the ability to attain a higher operating voltage with limited maximum device rating as well as fewer harmonics in the output current waveform. However, more semiconductor devices are needed and the control is more complicated. The main aim of this work is to present a method to predict the inverter lifetime. Therefore, the basic three-phase half-bridge inverter topology is selected for simplicity. PLECS is used as the circuit simulator since it runs fast, is simple to understand and supports user defined thermal description for the power devices.

4.3.1 Main Electrical and Thermal Circuit

A three-phase half-bridge inverter model is built in PLECS, which implements the electrical and thermal models introduced in the last two sections to obtain T_j form a particular mission profile. Its main electrical and thermal circuit is shown in Fig. 4.19. Each pair of IGBTs and free wheel diodes share a heatsink since they are geometrically located close to each other and are mounted on the same DCB. The parameters of each IGBT are exactly the same in the simulation, hence the junction temperatures of each IGBT are almost the same during operation. Consequently, the assumption is made that there is no heat flow between the each heatsink. The water temperature is assumed to be constant and is set to be the initial temperature of the system.

4.3.2 Gate Drive Control: Space Vector PWM

The space vector pulse width modulation (SVPWM) technique has been widely investigated since 1990s'. Compared with the carrier based sine-triangle PWM (SPWM), SVPWM gives a higher output voltage for the same DC bus voltage, better harmonic performance and

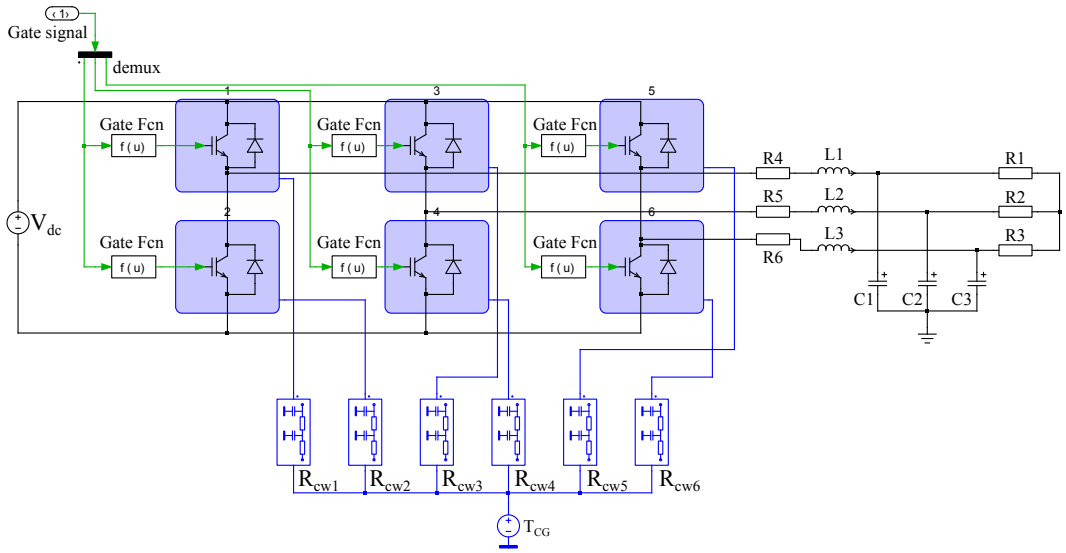


Figure 4.19: Main electrical and thermal circuit of a three phase half bridge inverter.

lower switching frequency which leads to lower switching loss [85, 86].

Fig. 4.20 shows the electrical circuit diagram for a three-phase voltage source PWM inverter to help explain the principles of SVPWM. S1 to S6 are six power IGBTs that are controlled by the switching variables a, a', b, b', c and c' respectively (1 is on and 0 is off). In order to avoid a short circuit in the DC bus, the lower IGBTs are switched off when the upper IGBTs are switched on and vice versa. Therefore a'=-a, b'=-b and c'=-c. The on and off states of the IGBTs can be used to determine the inverter output voltage according to Eqns. 4.10 and 4.11.

$$\begin{bmatrix} V_{ab} \\ V_{bc} \\ V_{ca} \end{bmatrix} = V_{dc} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{4.10}$$

$$\begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} = \frac{V_{dc}}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{4.11}$$

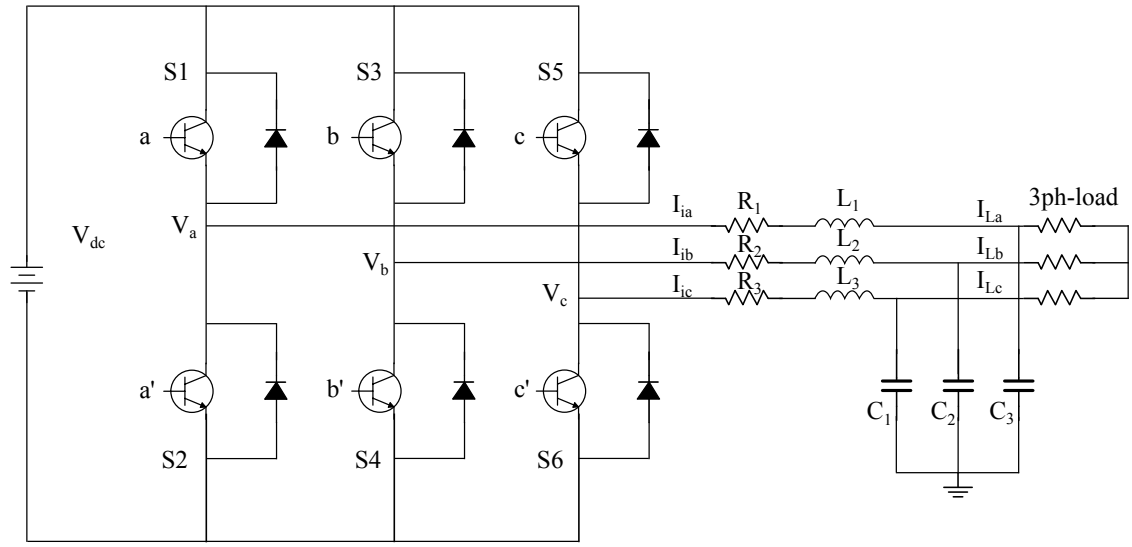


Figure 4.20: Main electrical circuit of three phase inverter.

From Eqn. 4.11, it is clear that the maximum phase voltage is $2V_{dc}/3$, while the maximum phase voltage of SPWM is only $V_{dc}/2$. There are eight possible combination of on and off patterns for the inverter controlled by SVPWM signal. The output line and phase voltages in terms of V_{dc} can be calculated from Eqns. 4.10 and 4.11 and are listed in Table 4.6.

To implement the SVPWM, a reference voltage vector V_{ref} is sampled at a sampling frequency f_s. The reference voltage vector can be obtained by either measurement or estimation of the three phase load voltage. By applying the abc-$\alpha\beta$ transformation, the reference volt-

Table 4.6: Switching vectors, line voltages and phase voltages in each voltage vector

Voltage vectors	**Switching Vectors**			**Line voltages**			**Phase voltages**		
	a	b	c	V_{ab}	V_{bc}	V_{ca}	V_a	V_b	V_c
V_0	0	0	0	0	0	0	0	0	0
V_1	1	0	0	1	0	-1	2/3	-1/3	-1/3
V_2	1	1	0	0	1	-1	1/3	1/3	-2/3
V_3	0	1	0	-1	1	0	-1/3	2/3	-1/3
V_4	0	1	1	-1	0	1	-2/3	1/3	1/3
V_5	0	0	1	0	-1	1	-1/3	-1/3	2/3
V_6	1	0	1	1	-1	0	1/3	-2/3	1/3
V_7	1	1	1	0	0	0	0	0	0

age vector can be described as a rotating vector on a two dimensional perpendicular frame at an angular frequency of $\omega = 2\pi f$, where f is the grid frequency. The equations in the abc reference frame can be transformed into the stationary α-β reference frame according to Eqn. 4.12.

$$\begin{bmatrix} f_\alpha \\ f_\beta \end{bmatrix} = \frac{2}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} f_a \\ f_b \\ f_c \end{bmatrix} \tag{4.12}$$

Where f denotes either voltage or current variable.

The objective of the SVPWM technique is to approximate the reference voltage vector V_{ref} using the eight voltage vectors. The eight voltage vectors on the stationary reference frame are shown in Fig. 4.21. The reference voltage in each sector can be described by its adjacent stationary voltage vectors according to Eqn. 4.13.

$$V_{ref} = \frac{T_k}{T_s} V_k + \frac{T_{k+1}}{T_s} V_{k+1} \tag{4.13}$$

Where, T_k and T_{k+1} are the dwelling time for V_k and V_{k+1} respectively, T_s is the switching period ($T_s = 1/f_s$), k is the sector that V_{ref} is in.

The reference angle θ, which is defined as the angle between the reference voltage vector V_{ref} and axis α, is used to determine the sector in which the reference voltage vector V_{ref} is located at a given time. θ can be calculated by Eqn. 4.14

$$\theta = tan^{-1}(\frac{V_{ref\beta}}{V_{ref\alpha}}) \tag{4.14}$$

However, θ calculated from Eqn. 4.14 always lies between $-\pi/2$ and $\pi/2$, which cannot represent the left side of the reference frame. Therefore, it is necessary to find a unique condition to represent each sector, as shown in Table. 4.7. Define V_{refA}, V_{refB} and V_{refC} by Eqns. 4.15, 4.16 and 4.17 respectively. If $V_{refA} > 0$, $A = 1$; otherwise $A = 0$. B and

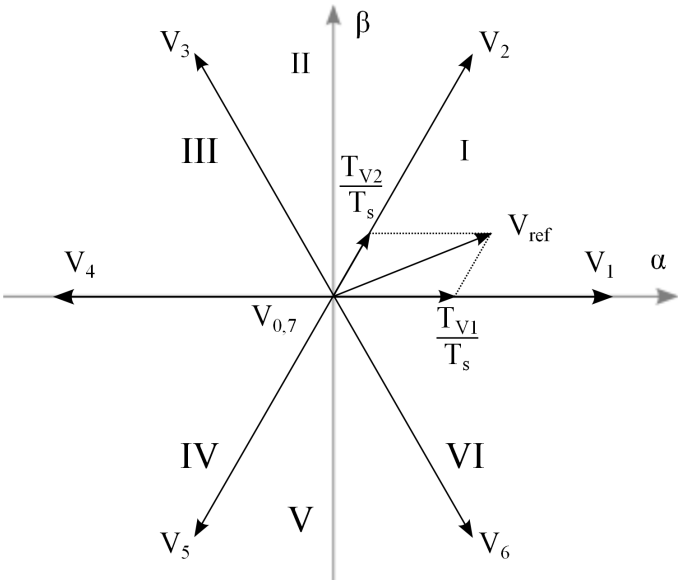


Figure 4.21: Basic switching vectors and sectors.

C follow the similar rule. N, which is one to one correspond with sector number k, can be calculated according to Eqn. 4.18 [87]. The sector selector model is built according to above definitions and is shown in Fig. 4.22.

Table 4.7: Necessary and sufficient conditions for each SVPWM sector.

Sector k	Conditions	A	B	C	N
1	$0 \leq V_{ref\beta} < \sqrt{3}V_{ref\alpha}$	1	1	0	3
2	$0 \leq \sqrt{3}V_{ref\alpha} \leq V_{ref\beta}$ or $0 < -\sqrt{3}V_{ref\alpha} < V_{ref\beta}$	1	0	0	1
3	$0 < V_{ref\beta} \leq -\sqrt{3}V_{ref\alpha}$	1	0	1	5
4	$\sqrt{3}V_{ref\alpha} < V_{ref\beta} \leq 0$	0	0	1	4
5	$V_{ref\beta} \leq \sqrt{3}V_{ref\alpha} \leq 0$ or $V_{ref\beta} < -\sqrt{3}V_{ref\alpha} < 0$	0	1	1	6
6	$-\sqrt{3}V_{ref\alpha} \leq V_{ref\beta} < 0$	0	1	0	2

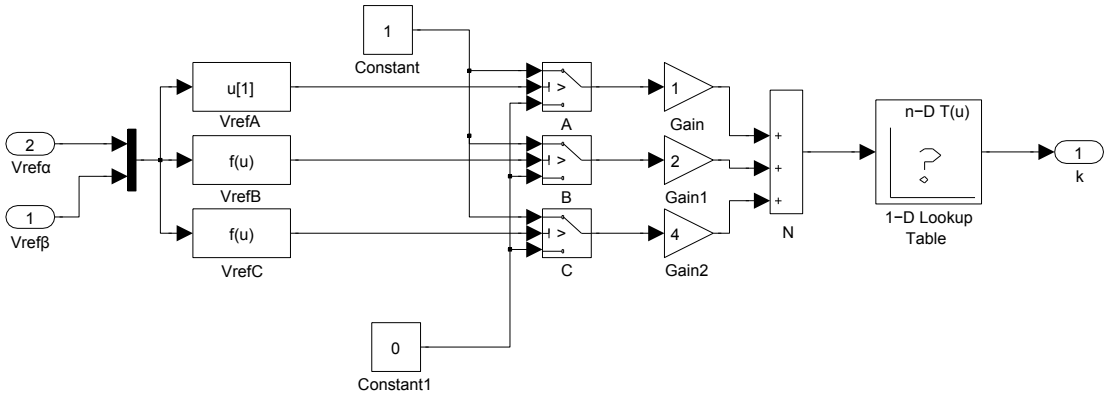


Figure 4.22: Sector selector model.

$$V_{refA} = V_{ref\beta} \tag{4.15}$$

$$V_{refB} = \sqrt{3}V_{ref\alpha} - V_{ref\beta} \tag{4.16}$$

$$V_{refC} = -\sqrt{3}V_{ref\alpha} - V_{ref\beta} \tag{4.17}$$

$$N = A + 2B + 4C \tag{4.18}$$

The aim of the modulation method is to control the switching sequence of the inverter to output the desired voltage or current. Therefore, one important step is to determine the dwelling time for each IGBT. They can be calculated with a known V_{ref} and k according to Fig. 4.21. In the case of $k = 1$, Eqns. 4.19 and 4.20 can be obtained.

$$|V_1|T_1 + |V_2|T_2 cos\frac{\pi}{3} = V_{ref\alpha}T_s \tag{4.19}$$

$$|V_2|T_2 sin\frac{\pi}{3} = V_{ref\beta}T_s \tag{4.20}$$

Knowing that $|V_1| = |V_2| = 2V_{dc}/3$. T_1, T_2 and T_0 can be calculated by solving Eqns. 4.19 and 4.20:

$$T_1 = \frac{3V_{ref\alpha} - \sqrt{3}V_{ref\beta}}{2V_{dc}}T_s \tag{4.21}$$

$$T_2 = \frac{\sqrt{3}V_{ref\beta}}{V_{dc}}T_s \tag{4.22}$$

$$T_0 = T_7 = \frac{T_s - T_1 - T_2}{2} \tag{4.23}$$

Eqns. 4.21, 4.22 and 4.23 are derived in sector 1 and can be expanded to the other five sectors. In order to simplify the description, X, Y and Z are defined by Eqns. 4.24, 4.25 and 4.26 respectively. Table 4.8 shows the dwelling time of the basic space voltage vector in each sector.

$$X = \frac{\sqrt{3}V_{ref\beta}}{V_{dc}}T_s \tag{4.24}$$

$$Y = \frac{\sqrt{3}V_{ref\alpha} + 3V_{ref\beta}}{2V_{dc}}T_s \tag{4.25}$$

$$Z = \frac{-\sqrt{3}V_{ref\alpha} + 3V_{ref\beta}}{2V_{dc}}T_s \tag{4.26}$$

Table 4.8: The dwelling time of basic space voltage vector in each sector.

Sector	**1**	**2**	**3**	**4**	**5**	**6**
T_k	-Z	Z	X	-X	-Y	Y
T_{k+1}	X	Y	-Y	Z	-Z	-Z

In the case of over modulation when $T_k + T_{k+1} > T_s$, the output waveforms will lead to major distortions [87]. Some modifications are needed:

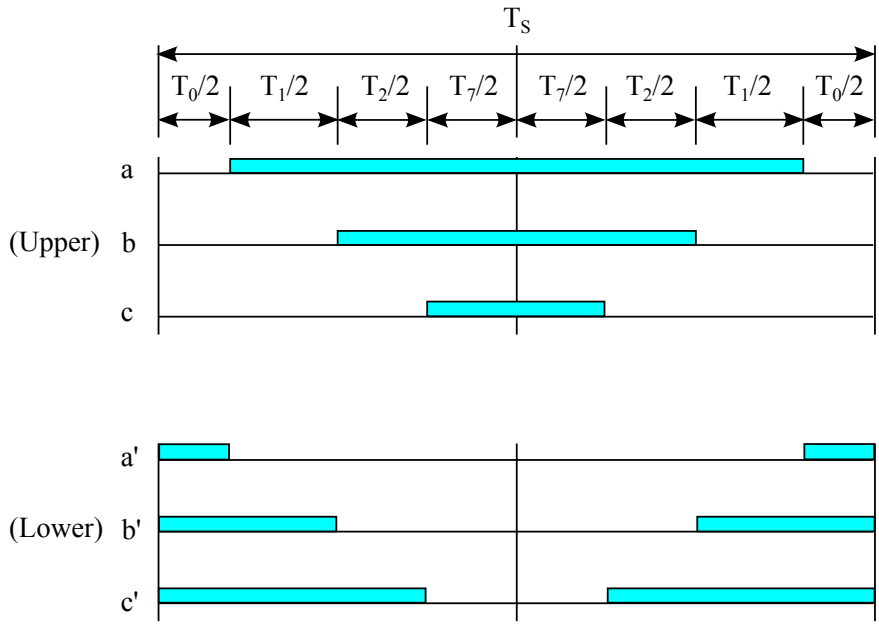


Figure 4.23: SVPWM switching patterns in sector 1.

$$T'_k = \frac{T_k}{T_k + T_{k+1}} T_s \tag{4.27}$$

$$T'_{k+1} = \frac{T_{k+1}}{T_k + T_{k+1}} T_s \tag{4.28}$$

There are three switch points in each sector: T_a, T_b and T_c. The switching time is arranged according to the first half of the switching period. The other half is simply the reflection of the first half. Fig. 4.23 shows the switching patterns in sector 1 for SVPWM modulation. In order to simplify the description, T'_a, T'_b and T'_c are defined by Eqns. 4.29, 4.30 and 4.31 respectively. Table 4.9 shows the switch time of each inverter leg in each sector. The structure of SVPWM generator is shown in Fig. 4.24.

$$T'_a = \frac{T_s - T_k - T_{k+1}}{4} \tag{4.29}$$

Table 4.9: IGBT switching time in each sector.

Sector	1	2	3	4	5	6
T_a	T'_a	T'_b	T'_c	T'_c	T'_b	T'_a
T_b	T'_b	T'_c	T'_a	T'_b	T'_c	T'_c
T_c	T'_c	T'_a	T'_b	T'_a	T'_a	T'_b

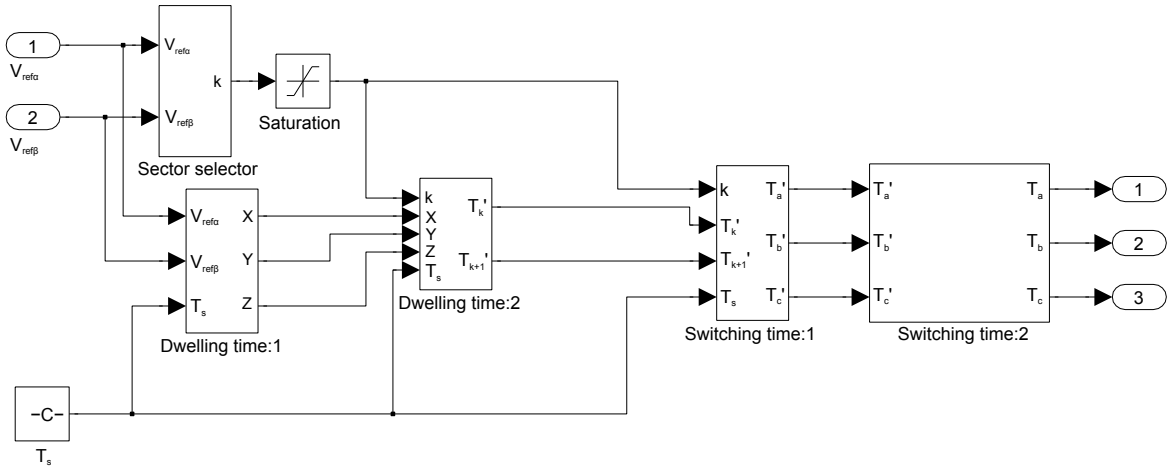


Figure 4.24: SVPWM generator model.

$$T'_b = T'_a + \frac{T_k}{2} \tag{4.30}$$

$$T'_c = T'_b + \frac{T_{k+1}}{2} \tag{4.31}$$

Chapter

5 Design of the Power Cycling Rig and Test

In actual applications, converters are not always working under constant load stresses. Hence T_j would fluctuate as the load stress varies. The temperature cycle and CTE mismatch between adjacent layers could cause a cyclic shear stress hence lead to the fatigue failures of IGBT modules. Therefore, T_j is a key parameter that affects the lifetimes of IGBT modules. As discussed in chapter 3, IGBT die-attach solder and bond wire suffer the highest thermomechanical stresses. Power cycling tests are designed in order to investigate how the cyclic temperatures affect the device lifetime and which failure mechanism is the dominant. The aim of such test is to develop a thermomechanical model to predict the device lifetime during inverter operation. Therefore, the requirements of this rig are listed below:

- Record the maximum and minimum of T_j/T_c of each IGBT during every cycle.
- Record the maximum V_{CE} of each IGBT during heating.
- Constant amplitude for each temperature cycle.
- Heating and cooling the IGBT modules as fast as possible.
- Degradation checks in each cycle to monitor the conditions of each IGBT.

- Automatic process control and data recording.
- Fault detection and protection throughout the test.

This chapter will introduce the design of both the power cycling rig and tests, as well as a brief result analysis.

5.1 System Configuration of Power Cycling Test

Power cycling refers to the act of repeatedly turning the device on and then off. An IGBT chip heats up by the power dissipated in it during the heating phase and cools down by the heatsink during the cooling phase, which composes a full temperature cycle. The maximum T_j can be controlled by adjusting the duration of heating period. Fig. 5.1 shows the picture of the power cycling rig. This rig is controlled by a computer and is capable of testing eleven half-bridge IGBT modules at the same time. This rig can be divided into three subsystems: the main electrical system, the water cooling system, and the control and data acquisition system.

5.1.1 Main Electrical System

The main electrical circuit of the power cycling test rig is shown in Fig. 5.2. Eleven half-bridge IGBT modules (SKM50GB123D) are connected in series so that the current flowing through each IGBT chip is the same, this ensures the temperature cycle of each module is in a similar range. The rated current, 50 A, is supplied to these modules during the heating phase to obtain the maximum possible heating efficiency. An IGBT module with higher rating (SKM75GB123D) is selected to work as a switch to control the on/off state of the main power supply since the switch experiences higher electrical stress. Its gate control signal is generated by the Labview program. An inductor is connected in the circuit to limit the current overshoot when the switch turns on. This circuit requires a power supply that can

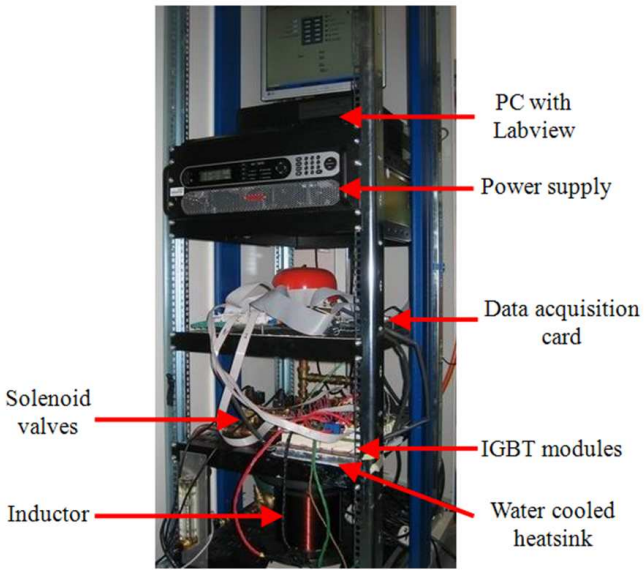


Figure 5.1: Power cycling rig designed to investigate the temperature effects on IGBT module lifetime

work as both voltage source (cooling phase) and current source (heating phase). Therefore, a programmable DC power supply (Sorensen SGI 100/100) is a reasonable choice. The gate voltage of the eleven IGBT test modules are kept at 15 V, since the modules are required to stay on throughout the test. A 100 mA current source is connected to all test modules to measure T_j indirectly [17]. The principle of the T_j measurement method will be discussed in the following section.

5.1.2 Water Cooling System

A water cooling system is designed to cool the IGBT modules as quickly as possible. Fig. 5.3 shows the configuration of the system. A chiller (Lauda WK 4600) is used to keep the cooling liquid at a constant temperature (15 ℃). All IGBT modules are mounted on a heatsink. A pump is connected to boost the water flow in order to cool the modules faster. The path of water flow is controlled by two solenoid valves. In order to shorten the period of power cycle, water flows through the bypass valve only during the heating phase. In the case

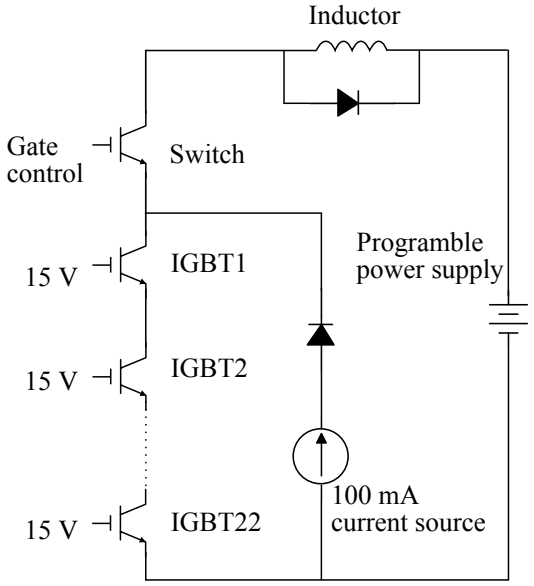


Figure 5.2: Main electrical circuit diagram of the power cycling rig.

of cooling phase, the low temperature coolant flows through the heatsink to cool down the IGBT modules. A time gap is applied between the operation of two valves to make sure at least one valve is turned on. The temperature of the water flow in and out the heatsink are measured and can be used as an alternative method to estimate the IGBT power dissipation.

5.1.3 Control and Data Acquisition System

The system is controlled by a computer with Labview installed. A data acquisition card (NI USB-6225 OEM) is selected to collect the experimental data and generate control signals. Current, water flow rate, T_{win}, T_{wout}, V_{CE}, T_j and T_c are recorded during each cycle. Detailed I/O list of the system is available in Appendix D.2.

The overview of the control algorithm of the power cycling test is shown in Fig. 5.4. The system first runs the initialisation and prepares for the test after it is started. Device degradation is monitored at the end of each heating phase by comparing the current V_{CE} and R_{th} with their initial values. The system turns into cooling phase if no IGBT is degraded. The counter which indicates current number of cycle is checked after the cooling phase. The

system starts another cycle and increases the counter by one until it reaches the desired number of cycle.

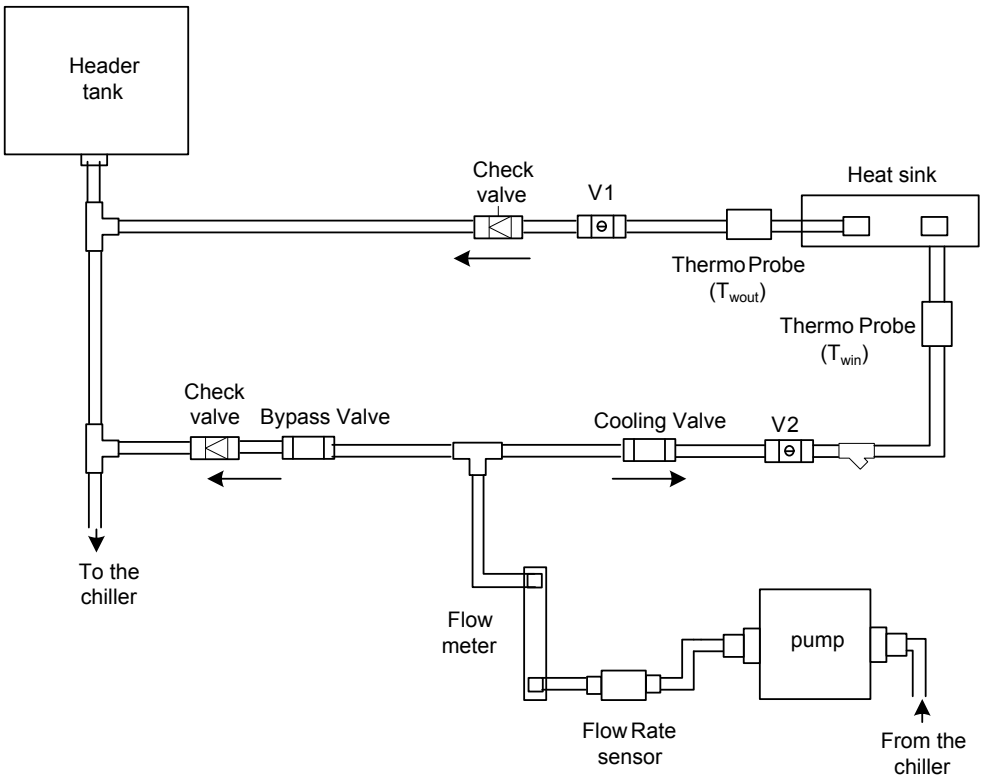


Figure 5.3: The configuration of water cooling system.

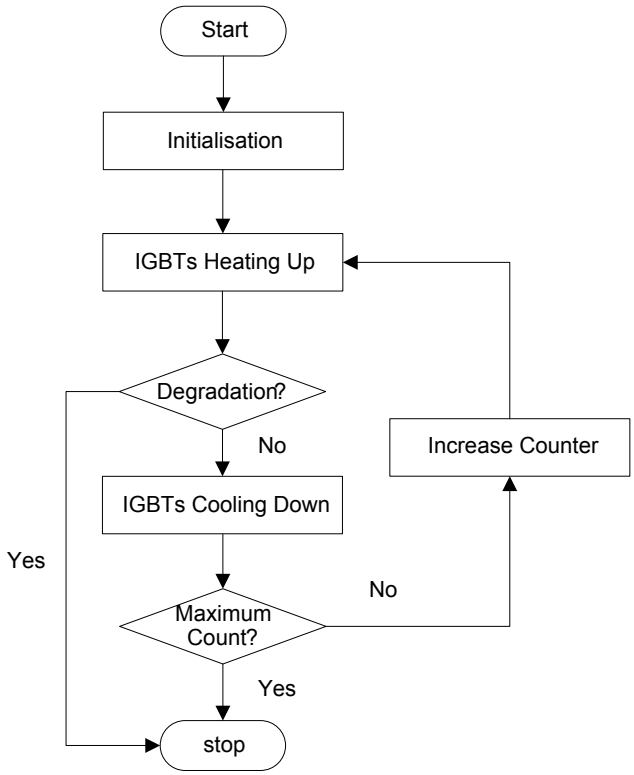


Figure 5.4: Simplified control algorithm of power cycling test.

The stop procedure is defined and listed below:

1. Turn-off the power supply and the gate drive of the switch; turn-on the cooling valve.
2. Time delay.
3. Turn-off the bypass valve.
4. Turn-on the relevant status LED and turn-off the healthy LED on the front panel.

There are several parameters that need to be set before the test, such as maximum number of cycles, degradation thresholds and maximum T_c. System initialisation, which includes 8 steps as shown in Fig. 5.5, is designed to read the preset parameters from the front panel and ensure the system is ready to run. The initial state of the system is defined as: the gate drive of the switch is off; the output of power supply is enabled; the bypass

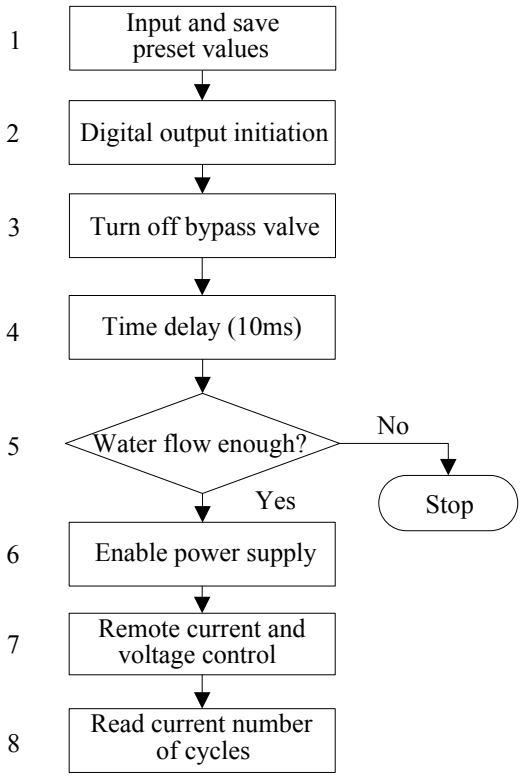


Figure 5.5: Detailed flow chart for initialisation.

valve is turned off; and the cooling valve is turned on. The water flow rate is checked after the bypass valve is turned off to guarantee efficient cooling. Power supply is then enabled with its voltage and current set to be the designed value (80 V and 50 A). After all these steps, the current number of cycles is loaded from a data file and the system is ready to run.

The detailed flow chart for the heating phase is shown in Fig. 5.6. In this phase, the bypass valve is turned on and the cooling valve is turned off, hence the coolant stops cooling the heatsink. The control switch is turned on and 50 A current flows through each IGBT to heat them up. Current number of cycle (CNC) is updated in the data file in case of losing data in accidents, such as power shunt or PC restart. All measured data is backed up every 1000 cycles. All parameters that are required to record are measured repeatedly until T_c of any IGBT reaches the set upper limit (T_{cH}).

Several measured parameters are checked to ensure the system is working correctly. The

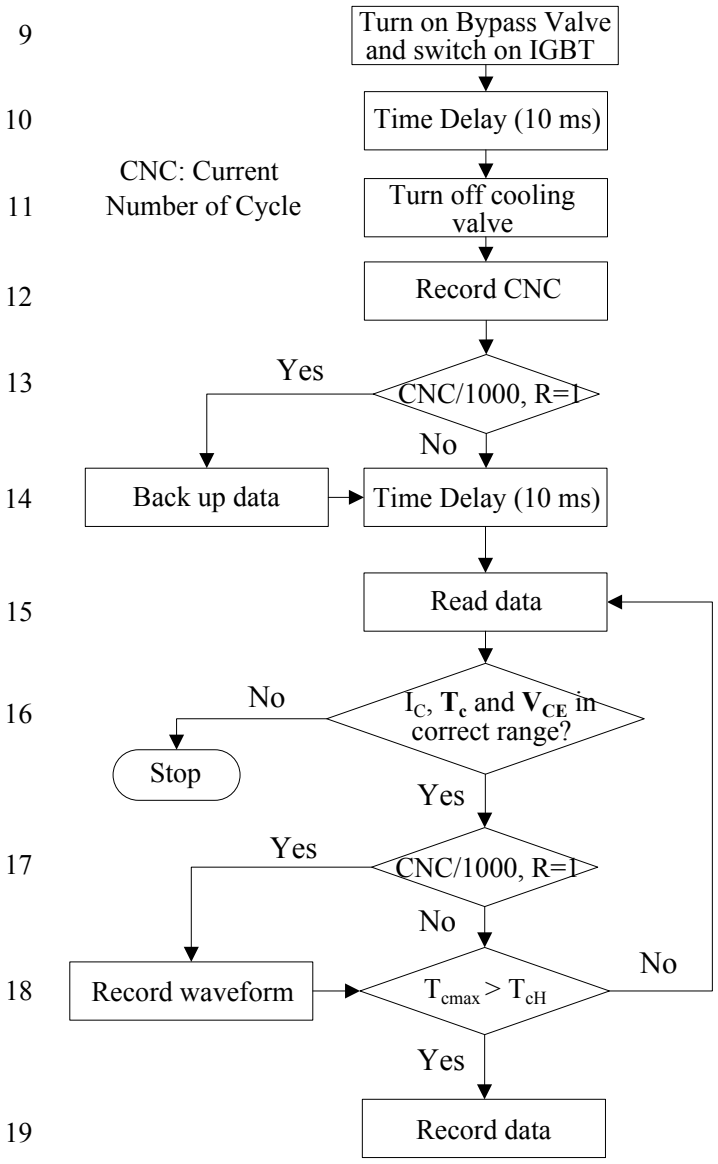


Figure 5.6: Detailed flow chart for IGBT heating phase.

Table 5.1: General diagnose of system failure by measured data

Parameter that are out of correct range	**System failure types**
$I_C = 0$ A	Open circuit failure
$V_{CE} < 1$ V	Short circuit failure
$T_c < 15$ ℃ or $T_c > 120$℃	Thermocouple failure

load current (I_C) should be maintained around 50 A, hence its upper I_{cH} and lower I_{cL} limit is set to be 55 A and 45 A respectively. T_c is an important parameter in the test, since it is used to control the duration of heating and cooling phase which determines ΔT_j. It is selected for two reasons. On the one hand, T_j can only be sensed when the main power supply is cut off, while T_c can be measured directly by thermocouples continuously. On the other hand, if T_j is selected, the increase of R_{th} would lead to a reduction of the dissipated energy, which is not realistic for real application conditions; while T_c excludes the influence of changes in cooling efficiency and includes the effects caused by the changes of thermal resistance of the module [57]. However as observed in the test, the thermocouple is not always reliable since it suffers the cyclic thermal stress as well. The two metal wires which are soldered together sometimes fall apart. Therefore, T_c monitoring is necessary to guarantee it is measured correctly. T_c for each IGBT are checked to maintain them in a reasonable range, between 15 ℃ and 120 ℃. V_{CE} is another important parameter which is used as the failure indicator of bond wire damage. V_{CE} of the selected IGBT module at 50 A within the operation is above 2 V. Therefore, any $V_{CE} < 1$ V normally indicates a short circuit failure. The system should be stop if that is detected. Several typical system failure examples and their indicators are listed in Table 5.1.

In the heating phase, the above parameters are measured and monitored continuously until it terminates when any T_c reaches I_{cH}. Only the last set of the sampling data which contains the maximum T_c and V_{CE} are recorded, unless the waveform recording is required.

T_j is a key parameter for the power cycling test. One main requirement of the rig is to maintain a constant ΔT_j for each cycle. Therefore, it is essential to measure the minimum and maximum T_j for each IGBT throughout the test. The minimum T_j appears at the end

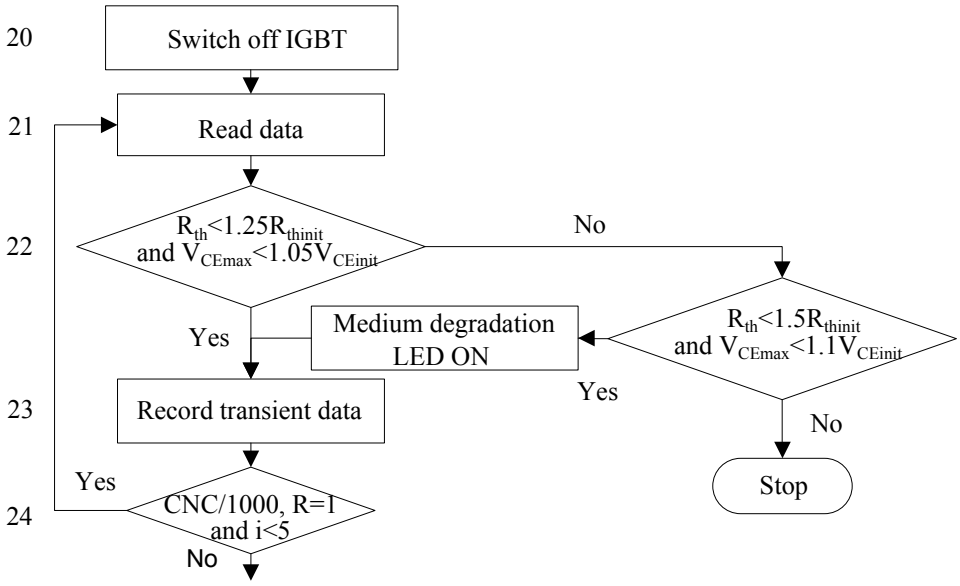


Figure 5.7: Detailed flow chart for IGBT degradation check.

of cooling phase while the maximum T_j appears at the end of heating phase. However, T_j can only be measured when the main power supply is cut off, hence the nearest possible measurement point appears right after the switch is turned off. As discussed in chapter 3, V_{CE} and R_{th} are the two failure indicators for bond wire damage and die-attach solder fatigue respectively. During the power cycling test, R_{th} for IGBT modules can only be estimated based on the assumption that the system is thermally stable at the end of heating phase, according to Eqn. 5.1. Therefore the degradation monitoring could only take place after V_{CE} and maximum T_j are measured, as shown in Fig. 5.7. The maximum sampling frequency of the DAQ card is 833 kHz and there are 80 channels available. Therefore, the maximum sampling frequency for each channel is 10.4 kHz, which means T_j can be measured after 0.1 ms of its theoretical maximum point. According to the thermal network of the IGBT module extracted in chapter 4, which short time constant is 0.1 s. Hence a measurement delay of 0.1 ms is too short for T_j to drop significantly. Therefore the measured T_j can be considered as the maximum value.

Table 5.2: Failure criteria for power cycling test

Type of degradations	Indicators
Die-attach solder medium degradation	R_{th} increase 25%
Die-attach solder serious degradation	R_{th} increase 50%
Bond wire medium degradation	V_{CE} increase 5%
Bond wire serious degradation	V_{CE} increase 10%

$$\frac{T_j - T_c}{P} = \frac{T_j - T_c}{V_{CE} \cdot I_C} = R_{th} \tag{5.1}$$

The system compares the current measured V_{CE} and estimated R_{th} with their initial values V_{CEinit} and R_{thinit} respectively. The initial value of a parameter is defined as its average measured value at the first 50 cycles. No IGBT is degraded if both the increase of V_{CE} and R_{th} are lower than the set medium degradation threshold. If any of them are greater than the medium degradation threshold but lower than the serious degradation threshold, the medium degradation LED will turn-on. If any of them exceeds the serious degradation threshold, the system will stop in order to protect the test modules from being damaged. The short thermal time constant for the IGBT module is 0.1 s, while the data is measured every 0.5 s during heating and cooling phase since a filter is applied for accuracy. Therefore, fast transient T_j measurement is desired to obtain a better waveform which is required every 1000 cycles. The failure criteria are listed in Table 5.2. Detailed discussion of this failure criteria is available in the last section of this chapter.

The system turns into the cooling phase after the degradation monitoring process. Fig. 5.8 shows the flow chart for the cooling phase. The bypass valve is turned off 10 ms after the cooling valve is turned on, thus coolant start flowing through the heatsink and cooling the IGBT modules. Similar to the heating phase, T_j and T_c for each IGBT are measured repeatedly to ensure they are in reasonable range until all T_c are lower than the set threshold T_{cL}. The water flow rate and V_{CE} over the switch is monitored as well to ensure the efficiency of cooling and the off-state of the main power supply. Only the last set of the sampling data, which contains the minimum T_j and T_c, is recorded unless waveform recording is required

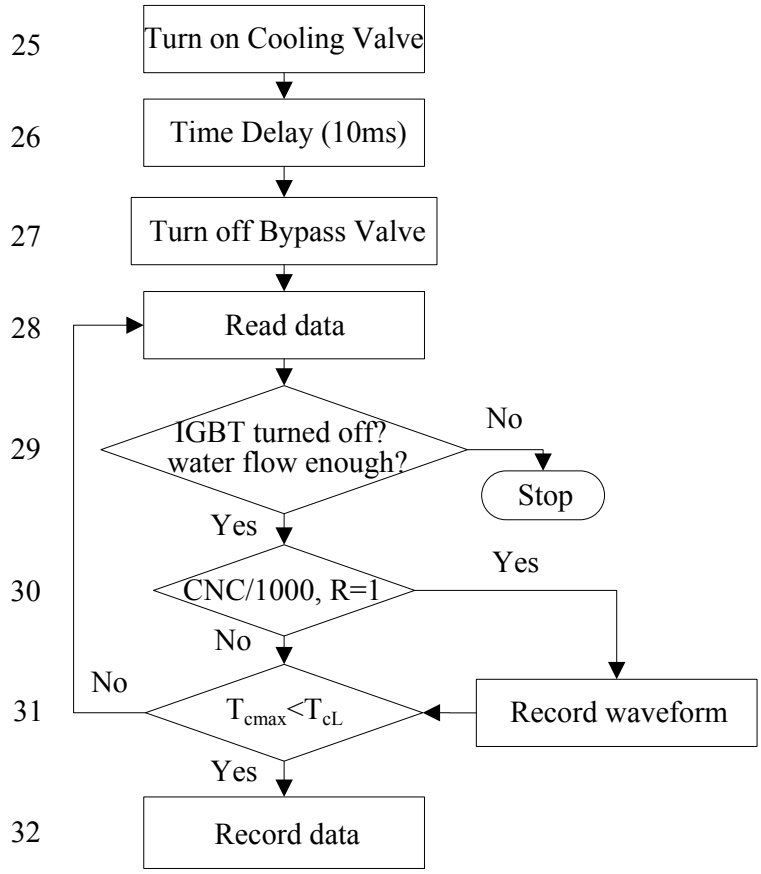


Figure 5.8: Detailed flow chart for IGBT cooling phase.

every 1000 cycles.

A protection circuit, dead mans handle (DMH), is designed to stop the power cycling and start cooling the heatsink if any error occurs in the control system, such as PC or DAQ card failure. A signal generation code is built and runs in parallel with the main program. This code generates a 1 Hz pulse signal to trigger a timer. The output of the timer turns to low if it is not triggered in two seconds from last trigger signal. The control signal of the cooling valve, the switch gate drive and the power supply are connected to an AND gate with the DMH signal respectively. The trigger signal will stop if any error occurs in the control system. Consequently, the cooling valve will turn-on; the switching IGBT will turn-off and the power supply will be disabled. This design protects the system based on hardware to

guarantee health and safety. The schematic for this circuit is available in Appendix D.3.

5.2 System Calibration and Preparation Before Test

After the system was built, it is important to calibrate the measurements to boost its accuracy. This section introduces the calibration methods and the preparations needed before the power cycling test.

5.2.1 Calibration of Junction Temperature Measurement

There are many different methods to measure T_j. They can be sorted into three categories: electrical, optical and physical contact [88]. Both optical and physical contact measurements can sense rapid variations in temperature and can easily provide temperature distribution of the chip surface. However, they require optical access to the chip which is not possible for packaged device.

Many electrical properties of the IGBT are highly temperature dependent. PN junction voltage drop, gate threshold voltage and leakage current are examples of temperature sensitive parameters (TSPs). A measurement of any above parameters can be used to estimate T_j. Although TSPs can only provide average temperature of the chip, it is widely used in real operation measurement since optical access is not available in most applications. In this work, PN junction voltage drop is selected as the TSP. T_j is estimated by measuring V_{CE} at a small calibration current (100 mA). According to Held [17], V_{CE} decreases linearly with a increasing T_j from 0 ℃ to 150 ℃ .

The calibration circuit is shown in Fig. 5.9. Three randomly selected test modules are connected in series and placed in the environmental chamber. The temperature of the chamber is controlled to increase from 25 ℃ to 150 ℃ while V_{CE} of each IGBT are recorded every 25 ℃. The test results are listed in Table 5.3, where $V_{CE1.1}$ and $V_{CE1.2}$ indicate the upper and lower IGBT of the first half-bridge IGBT module respectively. The linear relationship

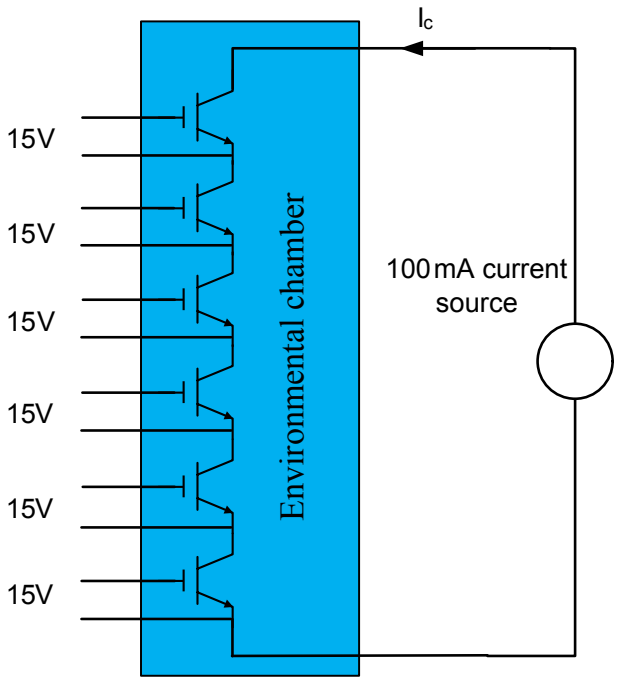


Figure 5.9: Test configuration of junction temperature measurement calibration.

between T_j and V_{CE} can be observed and is described by Eqn. 5.2. The fitting result is shown in Fig. 5.10. To obtain a more accurate T_j measurement, further calibration for each individual IGBT is designed before the power cycling test. T_j for each IGBT are measured at room temperature when the rig is off. Due to the individual differences, slight adjustment of Eqn. 5.2 is normally necessary to ensure all the measured T_j are equal to the room temperature.

$$T_j = -489V_{CE} + 307.2 \tag{5.2}$$

In order to build confidence on the TSP measurement method, infrared camera (IC) is used to measure T_j simultaneously. An IGBT module is opened and its silicone coat which covers the chip is removed, as shown in Fig. 5.11. This module is then mounted on the heatsink and connected to the test circuit used for the thermal network extraction, as shown in Fig. 4.13(a). Run the test program and measure T_j throughout the test by IC.

Table 5.3: Relationship between T_j (℃) and V_{CE} (V)

$\mathbf{T_j}$	$\mathbf{V_{CE1.1}}$	$\mathbf{V_{CE1.2}}$	$\mathbf{V_{CE2.1}}$	$\mathbf{V_{CE2.2}}$	$\mathbf{V_{CE3.1}}$	$\mathbf{V_{CE3.2}}$	Average
25	0.576	0.576	0.576	0.575	0.577	0.576	0.576
50	0.527	0.526	0.526	0.525	0.527	0.527	0.526
75	0.476	0.475	0.475	0.476	0.475	0.475	0.475
100	0.427	0.427	0.425	0.426	0.425	0.424	0.425
125	0.374	0.374	0.373	0.373	0.373	0.373	0.373
150	0.321	0.320	0.321	0.320	0.321	0.319	0.320

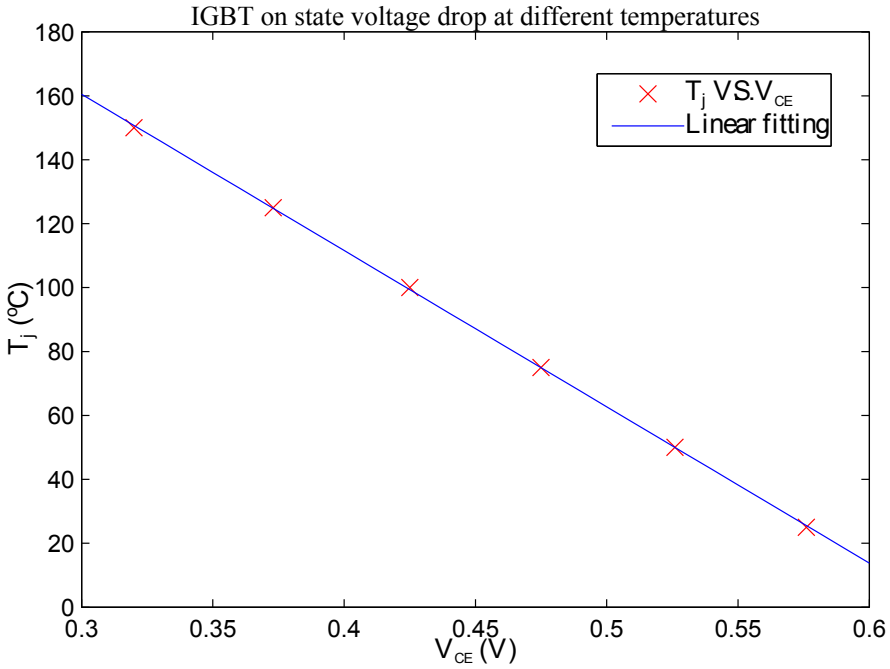


Figure 5.10: Calibration result of junction temperature measurement.

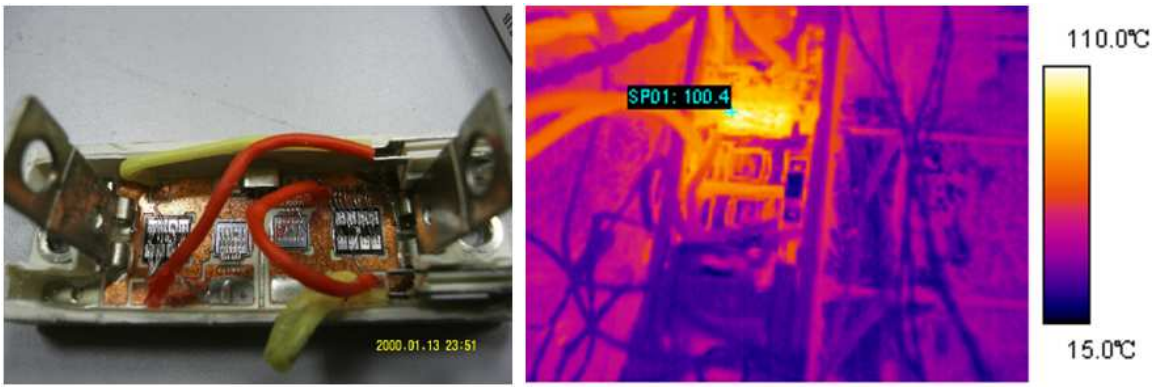


Figure 5.11: IGBT module under test.

The measured temperature curves are shown in Fig. 5.12. This TSP measurement method only sense the temperature at time point 2 as shown in Fig. 4.13(b) which is considered to be the maximum junction temperature of one load cycle. While the IC could measure the chip temperature in real-time, hence it recorded the temperature drop due to the cut-off of the main current as well. However, the TSP method can be considered as a temperature measurement method with low sampling frequency (same as the frequency of the cyclic current) since the temperatures measured by TSP and IC at the same time point are very close as shown in the figure. In this work, only the maximum and minimum T_j needs to be recorded hence real time measurement of T_j is not required. Therefore, the proposed TSP measurement method can be used and the accuracy meets the design requirement.

5.2.2 Calibration of Thermocouples

Any conductor subjects to a thermal gradient will generate a voltage, known as thermoeelctric effect. The main principle for thermocouple is using two different metals to complete a circuit in which two legs generate different voltages. The temperature difference between hot and cold junctions can be calculated by measuring the voltage difference at the cold junction, as shown in Fig. 5.13. To measure a single temperature, one junction, normally the cold junction, needs to maintain at a known constant reference temperature.

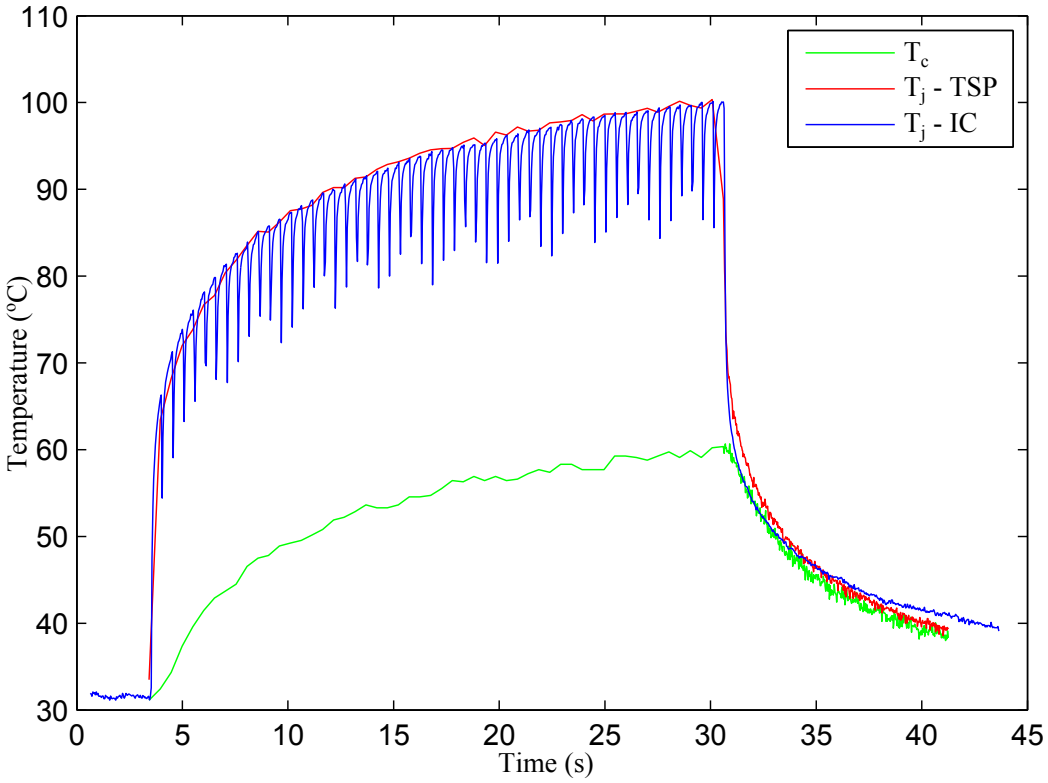


Figure 5.12: Comparing the T_j measured by TSP and IC.

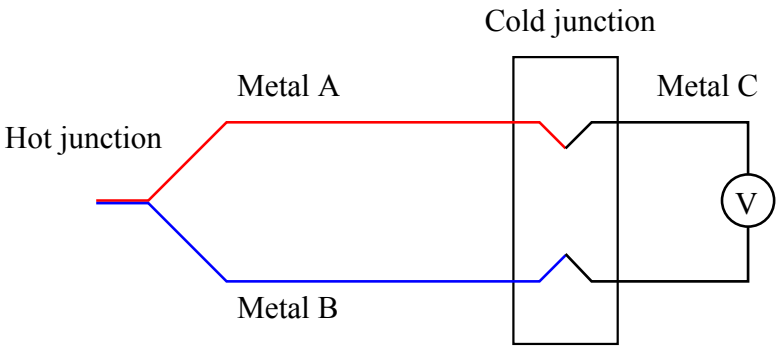


Figure 5.13: Principle of thermocouple.

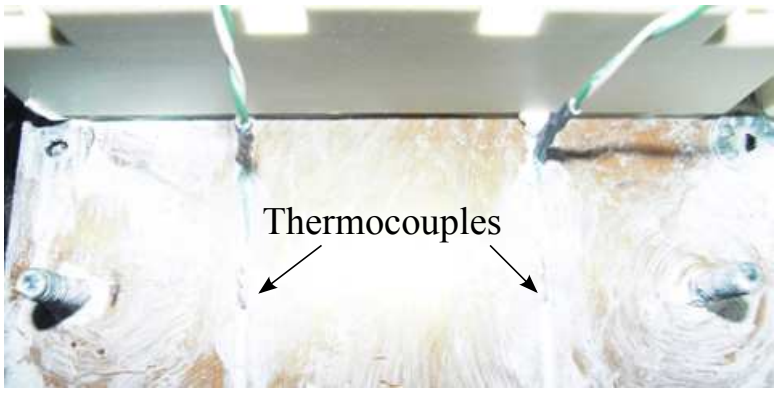


Figure 5.14: Thermocouple placement.

In this work, K type thermocouples are selected to measure T_c. This type of thermocouple is the most common general purpose thermocouple with a sensitivity of approximately 41 $\mu V/K$. It is able to sense temperature between -200 ℃ and +1350 ℃, which covers the designed T_c range. Four slots are trenched on the heatsink to bury the thermocouples. It is important to place the thermocouples in the same position right under each IGBT chip, as shown in Fig. 5.14. The misplacement of the thermocouple normally leads to a greater measured R_{th}, since the measured T_c will be smaller than the actual T_c. The measured T_c can be adjusted in Labview by changing the CJC (cold junction compensation) value. T_c measurement is calibrated by adjusting all measured T_c to the room temperature when the rig is off.

5.3 Design of Power Cycling Test and Data Processing

The degradation of the IGBT module packaging is caused by thermomechanical stresses. Therefore, T_j history is very important to IGBT module lifetime estimation. Some papers [4, 72] suggest that the number of cycles to failure (N_f) is affected by the parameters of the temperature cycle such as amplitude, mean temperature, frequency and dwell time. Although comprehensive multi-parameter models might be more accurate, they required a

more detailed parameter extraction process. Due to the time limitation, two sets of power cycling test are designed to study the effects of temperature cycle amplitude on the IGBT module lifetime. The mean temperature of these two tests are kept constant to exclude the disturb from different mean temperature. The maximum operation temperature of this IGBT module is 150 ℃ according to the datasheet in Appendix B, hence the designed test condition is in the safety operation zone. The test conditions are listed in Table 5.4.

Table 5.4: The designed test conditions

	T_{jmin}	T_{jmax}	ΔT_j	T_m	T_{cmax}
Test1	50 ℃	130 ℃	80 ℃	90 ℃	110 ℃
Test2	30 ℃	150 ℃	120 ℃	90 ℃	115 ℃

In each set of test, 22 IGBTs are connected in series and mounted on a water cooled heatsink. They were simultaneously heated by the rated 50 A current. After reaching the defined maximum T_c, load current is turned off and water cooling is turned on. The failure criteria are defined and have been shown in Table 5.2. These failure criteria have been defined, based on both previous researches [17, 57] and experimental observations. Held [17] proposed that 5% increase of V_{CE} indicates the bond wire damage. However due to the measurement errors shown in Fig. 5.16 and 5.17, 5% increase could be easily achieved. R_{th} is difficult to extract accurately during power cycling process. However, it can be estimated by the junction-case temperature difference and the IGBT power loss at the end of warming phase by assuming thermal stability at that time. Scheuermann [57] suggested that a 20% increase of R_{th} indicates the solder fatigue failure and used the maximum T_j directly to indicate R_{th}. However, the maximum T_j can be affected by the variation of either V_{CE} or the maximum T_c. The presented method removes all these effects and can be used when the warming phase period is relatively long. According to the experimental data, 50% increase of R_{th} is more representative to be the inflection point where R_{th} starts to increase significantly.

Furthermore, as the die-attach solder fatigue accumulates, the maximum T_j will increase and lead to the increase of V_{CE}. This means the rise of V_{CE} not only indicates the bond wire

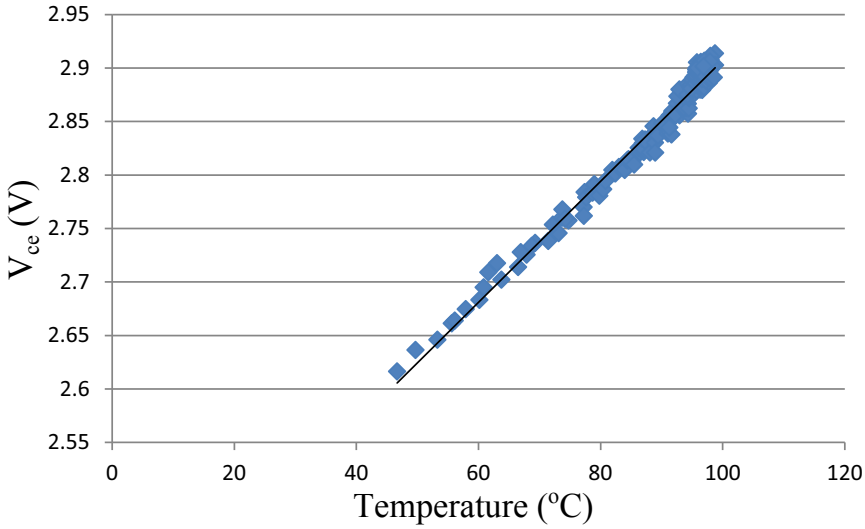


Figure 5.15: Calibration of V_{CE} versus T_j at 50 A.

damage but also the solder fatigue, which is not acceptable. To compensate the temperature effect on V_{CE}, it is desirable to understand how V_{CE} varies with temperature at the load current. The thermal network extraction test introduced in chapter 4 could also be used to calibrate V_{CE} versus T_j since it measures both parameters at the same time. Fig. 5.15 presents an example of the measured V_{CE} versus T_j. It is clear that V_{CE} increases linearly with the increase of T_j at 50 A, which is opposite when the current is 100 mA. Six IGBTs are randomly selected and Eqn. 5.3 can be fitted from the average of their test results. Therefore, the compensated V_{CE} can be calculated according to Eqn. 5.4.

$$V_{CE} = 0.0056T_j + 2.37 \tag{5.3}$$

$$V_{CE\_comp} = V_{CE} - 0.0056(T_{jmax} - T_{jmax\_initial}) \tag{5.4}$$

Here, T_{jmax} is the maximum T_j of one IGBT during heating phase and $T_{jmax\_initial}$ is defined as the average T_{jmax} of the first 50 cycles.

Fig. 5.16 shows the comparison between V_{CE} measurement with and without temperature

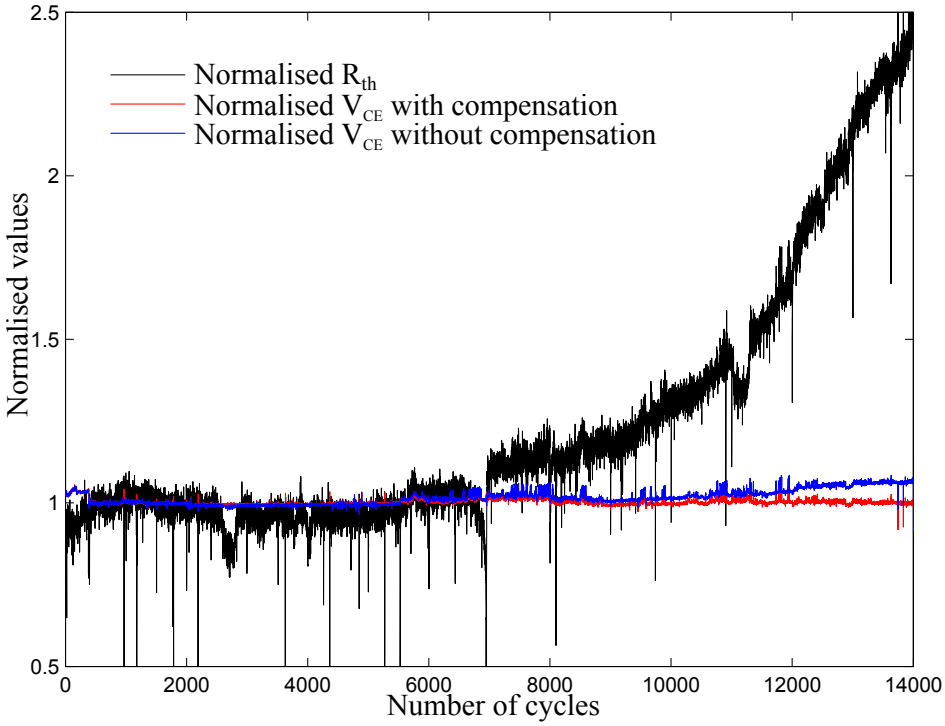


Figure 5.16: Comparison of V_{CE} measurement with and without temperature compensation.

compensation as well as the R_{th} degradation. Normalised values of R_{th} and V_{CE}, which are defined according to Eqn. 5.5, are used in order to keep the consistency. Here, f donates either R_{th} or V_{CE}, f_n is the normalised value, f_c is the current value and f_i is the initial value

$$f_n = \frac{f_c}{f_i} \tag{5.5}$$

It is clear that the measured V_{CE} is the same as V_{CE_comp} before the solder degrades. However, as the solder degradation accumulates, R_{th} starts to increase and this leads to the increase of T_j. The measured V_{CE} without temperature compensation starts to rise and tends to indicate that bond wire damage occurs, which is not true. While V_{CE_comp} keeps independent from the temperature variation and therefore can be considered as the unique

indicator of the bond wire damage failure.

Fig. 5.17 shows the typical results of the power cycling test. Clearly, solder degrades faster when ΔT_j is greater, which agrees with previous researchers [4, 57]. Some previous researchers proposed that bond wire damage is the main failure mechanism [17]. However in this designed test, the bond wire damage is only observed after die-attach solder is seriously degraded and junction temperature is above 200 ℃. This indicates that the solder fatigue is the dominant failure mechanism in the designed test conditions for the selected IGBT modules. However, the solder fatigue will not kill one device directly, the end of life failure is still bond wire damage.

The solder fatigue damage can be divided into two stages: crack initiation stage and crack propagation stage. During the crack initiation phase, cracks do not exist or are too small to cause damage. Therefore, the damage accumulation in this phase is nearly constant and can be thought of as being equal to zero. After a certain number of stress cycles, the crack develops to a critical length and starts to propagate, which represents the damage accumulation. From the experimental results shown in Fig. 5.17, it is the crack initiation stage which is highly ΔT_j dependent. While in the second stage, the cracks propagate at almost the same rate for both sets of tests, which is almost independent of ΔT_j.

The power cycling test results can be extracted from the recorded data according to the defined failure criteria and are listed in Table 5.5. Fig. 5.18 shows the summary of the power cycling results. Clearly, IGBT module lifetime decreases when ΔT_j rises. Furthermore, ΔT_j for each IGBT is different even in the same set of test. For test 1, the variation is 87 ± 7 ℃, while for test 2 it is 123 ± 13 ℃. This variation is caused by the individual difference between each IGBT. The differences on both on-state voltage drop or thermal resistance could lead to a quite different thermal behavior.

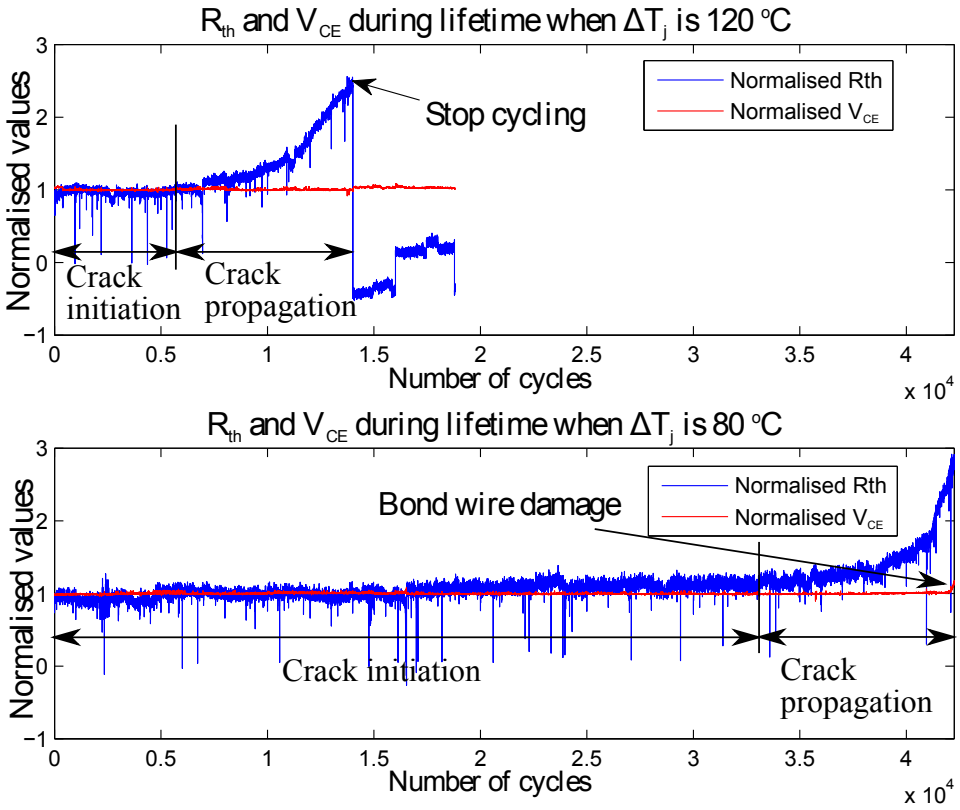


Figure 5.17: Typical results of power cycling test1 (lower) and test2 (upper).

Table 5.5: Power cycling results.

Power cycling test 1			**Power cycling test 2**		
IGBT Number	ΔT_j (°C)	N_f	**IGBT Number**	ΔT_j (°C)	N_f
1.1	94.3	39332	2.1	138.2	6286
1.2	93.1	44638	2.2	130.8	12931
1.3	92.8	53509	2.3	130.5	11429
1.4	91.4	51755	2.4	120.4	14842
1.5	89.2	51647	2.5	122.2	9948
1.6	92.6	42056	2.6	119.8	18770
1.7	91.1	48042	2.7	114.3	10838
1.8	84.8	62007	2.8	114.9	14937
1.9	82.5	60886	2.9	116.6	13658
1.10	84.9	60800	2.10	113.1	12223
1.11	84.3	N/A	2.11	123.4	11774
1.12	84.2	62021	2.12	126.3	9234
1.13	86.4	54219	2.13	121.1	11303
1.14	88.0	56808	2.14	122.7	9595
1.15	84.3	62643	2.15	120.4	11954
1.16	84.9	57396	2.16	124.1	12506
1.17	87.7	52834	2.17	124.5	9812
1.18	80.1	N/A	2.18	124.5	16438
1.19	81.7	62848	2.19	113.9	16324
1.20	81.6	51532	2.20	110.7	14892

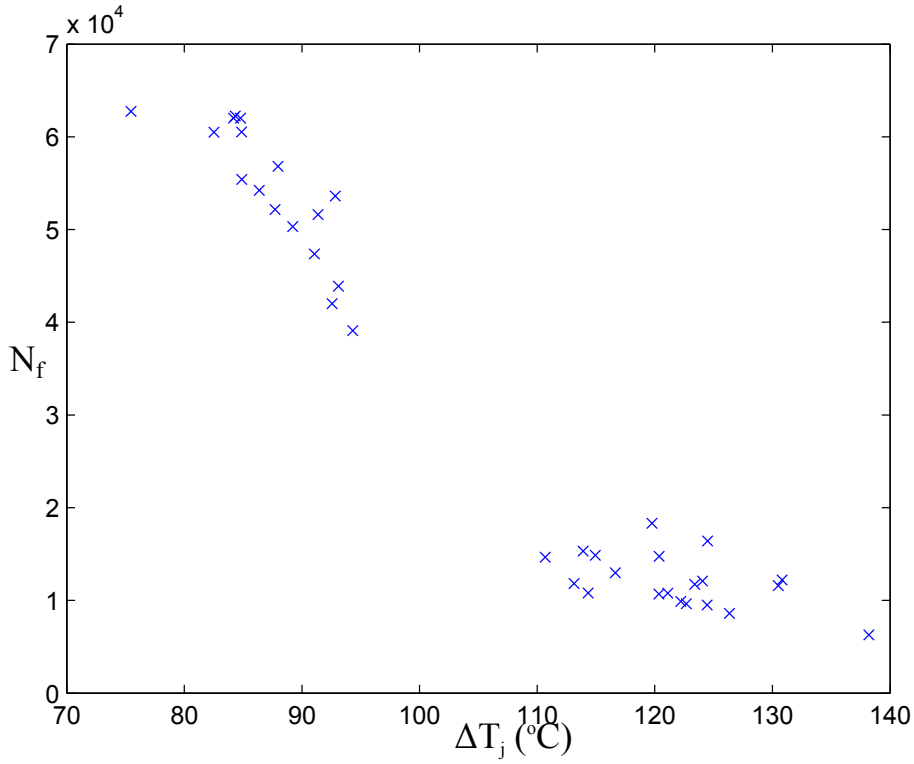


Figure 5.18: Lifetime data of all IGBTs with different ΔT_j.

5.4 Conclusions

This chapter explained the detailed design of the power cycling rig as well as the power cycling test. The failure criteria is defined based on both the literature and the experimental data. The junction temperature measurement using V_{CE} at 100 mA is calibrated and then proved by IC measurement result. V_{CE} at on-state increases with temperature, hence a compensation is carried out to remove the temperature effect and make it a necessary and sufficient indicator of bond wire damage.

The power cycling results are recorded and briefly analyzed in this chapter. From these results, it is clear that greater ΔT_j will lead to a shorter lifetime. However, only the crack initiation is greatly temperature dependent. While in the crack propagation stage, damage caused by different ΔT_j appear to be similar, which means independent with temperature cycle.

Chapter

6 Inverter lifetime Damage Model

This chapter presents the implementation of the inverter lifetime damage model. Initially, the power module package failure models introduced in chapter 3 are fitted to the power cycling test results. The Weibull distribution is used to extract the reliability information for each set of power cycling test. The empirical failure models could only estimate the lifetime for a constant temperature cycle. However in the case of a real application, the temperature cycle normally varies with time. Therefore, a cycle counting method is needed to decompose the complex temperature profile and extract the number of cycles as well as their corresponding amplitudes in such a profile. Furthermore, a method is required to accumulate the damage caused by each counted cycle. The current available cycle counting methods are briefly reviewed in this chapter and the suitable method is selected. Finally, two damage accumulation methods are discussed and compared.

6.1 Thermomechanical models of the IGBT module

According to the power cycling test results, die-attach solder fatigue is the dominant failure mechanism and bond wire damage only occurs after the solder is seriously degraded which

lead to a high junction temperature. Therefore, only the parameters of the die-attach solder fatigue model can be extracted for the test data.

Statistical distributions which represent the scatter in product life are widely used to analyse the product lifetime data [12]. There are several commonly used lifetime distributions, such as exponential, normal, lognormal, Weibull and extreme value distributions. In the case of electronic devices, the Weibull distribution is preferred because it is able to model either increasing or decreasing failure rates simply. Eqn. 6.1 expresses the Weibull probability density ($f(x)$), where α is the scale parameter and β is the shape parameter. Fig. 6.1 presents the variation of Weibull probability density by changing its parameters. It is clear that β controls the slope of the distribution, the greater β is the sharper its slope; while α controls the scale of the distribution, the greater α is the wider its shape. Furthermore, the peak moves towards α with the increase of β. For $x < \alpha$, the failure probability is dominated by the polynomial function of x. While for $x > \alpha$, its shape is donimated by the exponential decay. The Weibull accumulative distribution ($F(X)$) is derived by integrating the Weibull probability density , as shown in Eqn. 6.2.

$$f(x) = \frac{\beta}{\alpha^{\beta}} x^{\beta-1} exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \tag{6.1}$$

$$F(x) = 1 - exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \tag{6.2}$$

The Weibull accumulative distribution needs to be fitted to the power cycling results listed in Table 5.5. In order to simplify the fitting process, Eqn. 6.2 is modified to the form of Eqn. 6.3, where $\alpha' = -1/\alpha^{\beta}$. By processing this modification, the complex exponential term is removed, leaving only the polynomial term. The fitting results for both set of tests are shown in Fig. 6.2. The Weibull parameters can be extracted from the fitting results. For power cycling test 1, $\alpha_1 = 59404$, $\beta_1 = 7.842$ while $\alpha_2 = 13196$, $\beta_2 = 4.482$ for test 2. The Weibull accumulative and probability density distribution for both tests are plotted

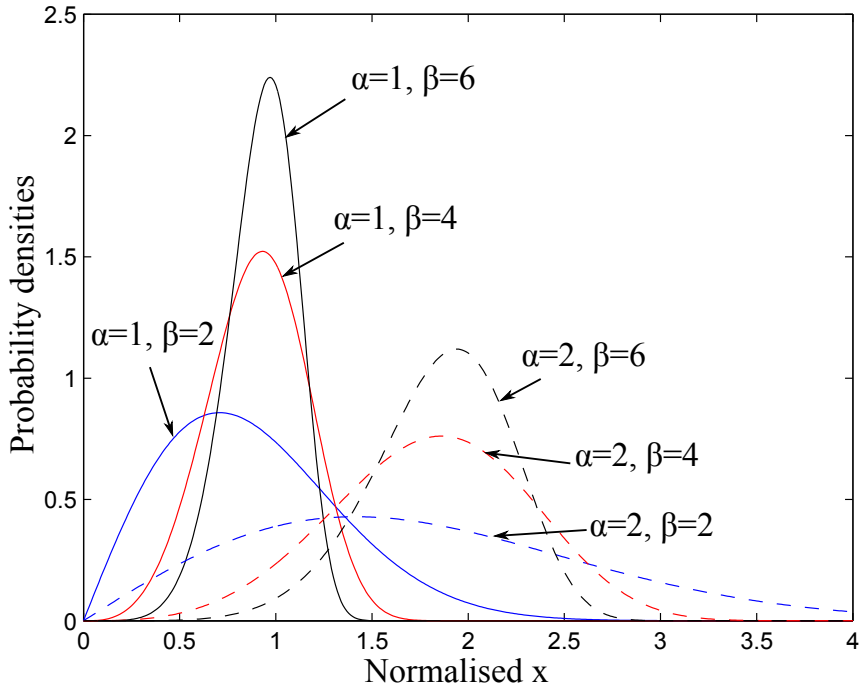


Figure 6.1: Weibull probability density distribution variation due to the change of parameters.

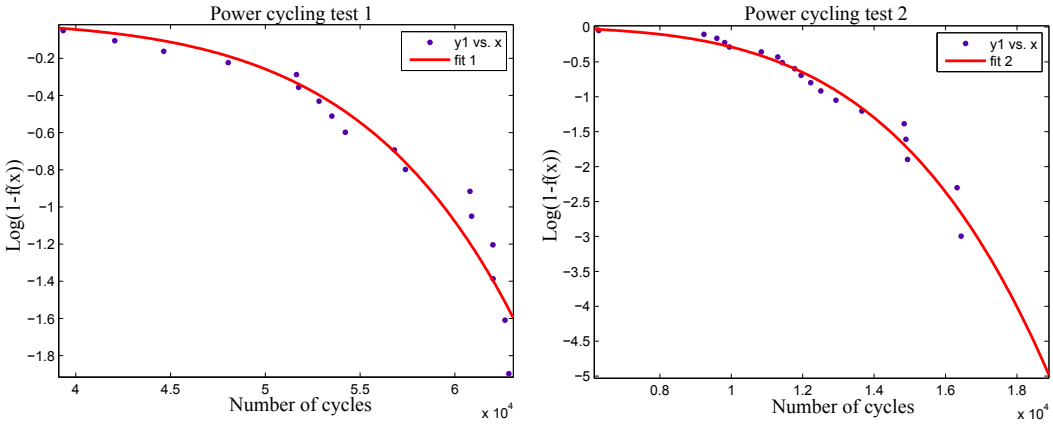


Figure 6.2: Fitting results for the modified Weibull accumulative distribution.

together with the power cycling results in Fig. 6.3 which shows a good representation of the lifetime scatter.

$$y = log(1 - F(x)) = -\left(\frac{x}{\alpha}\right)^{\beta} = \alpha' x^{\beta} \tag{6.3}$$

With the known Weibull parameters, the Weibull percentile can be calculated according to Eqn. 6.4. Therefore, the lifetime of 1% accumulative failures for both tests are 33042 and 4729 cycles respectively; while 50% accumulative failures are 56692 and 12160 cycles respectively.

$$N_P = \alpha[-ln(1 - P)]^{\frac{1}{\beta}} \tag{6.4}$$

The lifetime model for solder fatigue is given in chapter 3 by Eqn. 3.10. For simplification, it can be reduced to Eqn. 6.5 since the material properties and the power module structural dimensions are constants. As ΔT_j for each tested IGBT is different even within the same set of test due to individual device differences, their average value is used to represent the temperature cycle of the test. Thus $\Delta T_{j1} = 87$ ℃ and $\Delta T_{j2} = 123$ ℃. Substituting ΔT_j and N_f into Eqn. 6.5, the parameters for the fatigue lifetime model can be extracted, as listed in Table 6.1.

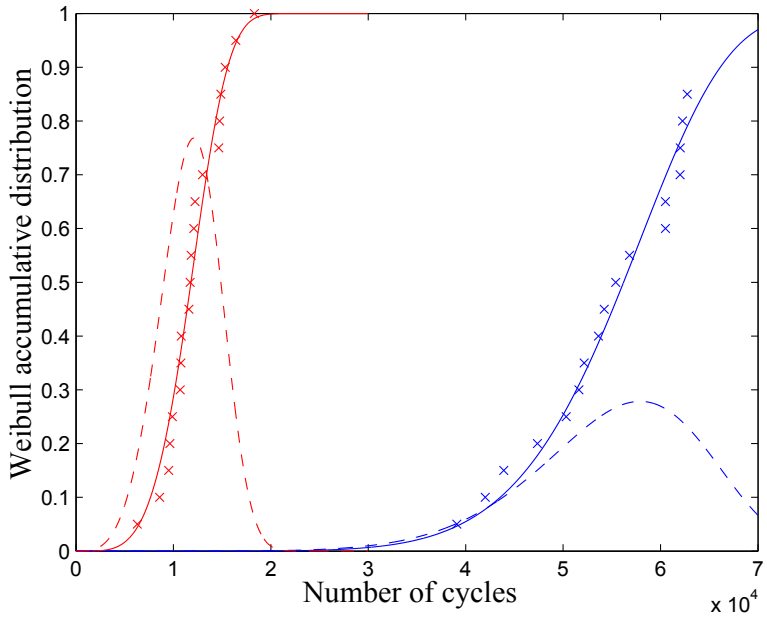


Figure 6.3: Fitted results for the Weibull accumulative and probability density distribution, blue indicates the results for power cycling test 1 and red indicates the results for power cycling test 2.

Table 6.1: Die-attach solder fatigue lifetime model for 1% and 50% Weibull accumulative failure.

	a'	**b'**	**Lifetime expression**
Weibull 1%	2.291×10^{15}	-5.6145	$N_f = 2.291 \times 10^{15}\Delta T_j^{-5.6145}$
Weibull 50%	2.377×10^{13}	-4.4457	$N_f = 2.377 \times 10^{13}\Delta T_j^{-4.4457}$

$$N_f = a'\Delta T_j^{b'} \tag{6.5}$$

Fig. 6.4 shows the lifetime extracted from 1% and 50% accumulative failure from Weibull analysis together with the power cycling results and an extrapolation of the LESIT results reported by Held [17]. Obviously, the power cycling results presented in this work have shown that the tested power modules have better reliability than those tested in the LESIT project for a wide range of ΔT_j. Furthermore, the lifetime model based on Weibull 50% accumulative failure shows a better fit with the power cycling results. The Weibull 1% fitting shows a more conservative lifetime estimation when $\Delta T_j > 55$ ℃ and a less conservative estimation for smaller ΔT_j. The same problem can be observed from the power cycling results reported by Scheuerman [57]. This is because the lifetime scatter under higher ΔT_j is wider than that under smaller ΔT_j. Hence for normal operation when ΔT_j is not as large as the accelerated test, the Weibull 50% results are used to obtain a more conservative estimation.

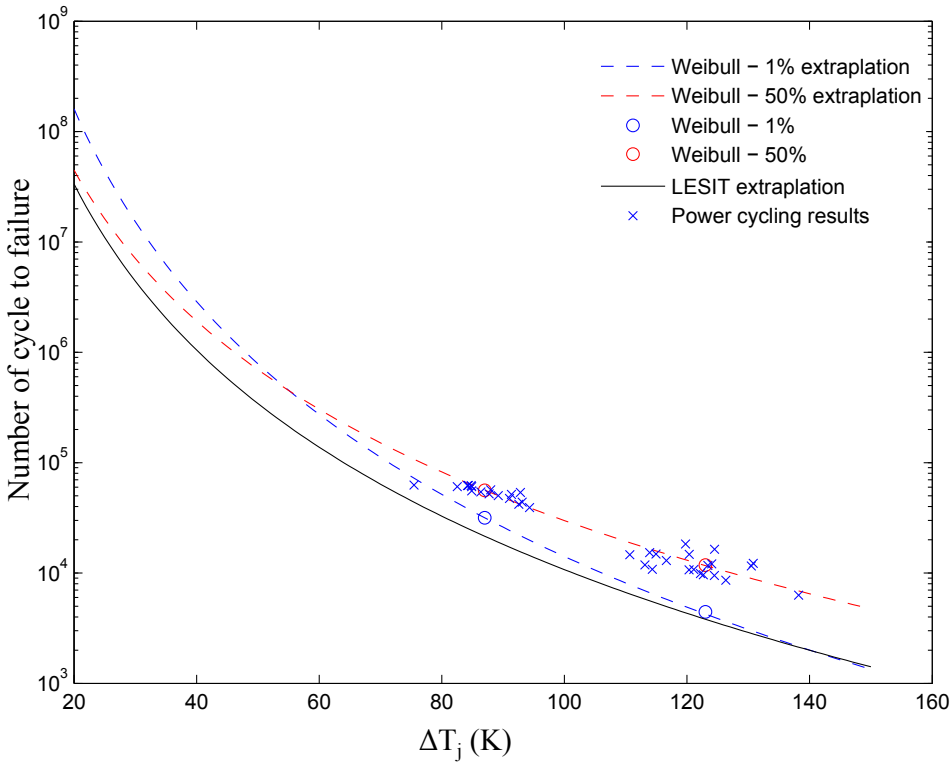


Figure 6.4: Comparison of the power cycling results with LESIT results.

6.2 Cycle counting method

There are several methods available to count irregular stress-versus-time profile, example included the level crossing counting method, peak counting method, simple range counting method and rainflow counting method. The definition of a cycle is different for each method. For level crossing method, the range between the maximum and minimum amplitudes of the stress profile is divided into discrete interval levels above and below the mean stress. A count is recorded whenever the stress profile crosses a present level above or equal to the mean stress with positive slope or a level below the mean stress with negative slope. Restrictions are needed to avoid small variations around the level to be counted which will lead to a large number of counts. This can be done by filtering the small noise prior to cycle counting or by ignoring the crossing until the adjacent next level is crossed. The counts are then combined to form completed cycles by first constructing the largest possible cycle within exist counts until all counts are used. Obviously, this method only gives the statistical summary of the stress amplitudes and consequently does not represent the actual stress profile.

The peak counting method records all the peaks above the mean stress and valleys below the mean stress. In order to eliminate negligible amplitudes, mean crossing peak counting is often applied, which counting only the largest peak or valley between two mean crossings. However, this restriction might eliminate some big cycles. The peaks and valleys are used to derive completed cycles according to the largest possible cycle rule, similar to the level crossing method. Therefore this method has the same shortcoming as the previous one, which is not representive to the actual stress profile.

A simple range counting method records the difference between the adjacent peak (above the mean) and valley (below the mean). The range from a peak to mean is termed negative range and its reversal is termed positive range. Each range is termed half cycle. Positive and negative ranges with the same amplitude can be paired to form a full cycle. This method is not recommended since it may miss out large cycles that do not cross the mean stress.

Rainflow cycle counting method, which was developed by Endo and Matsuishi in 1968

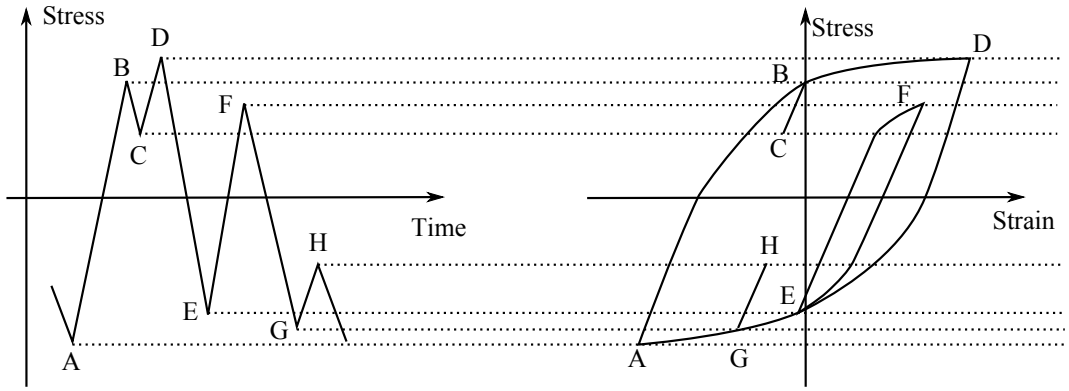


Figure 6.5: Rainflow cycle counting method based on hysteresis loops.

[89], is one of the most popular cycle counting techniques used in fatigue analysis. Cycles counted by this technique correspond to the fully closed hysteresis loops in the temperature profile, which is the physical basis of this method. Fig. 6.5 shows a stress profile and the hysteresis loops of the stress. As the material deforms from point A to D, it follows the path as shown in the hysteresis loop. For small stress cycles such as B to C, the stress is released slightly at point B and the material elastically shrinks to point C. When the stress increases from C to D, the material first remembers its prior history and elastically deforms to point B. After passing point B, the plastic deformation continues along path A to D as if the stress release from B to C never occurred. For large stress cycles such as E to G where the stress is reversed, plastic deformation occurs and appears as a small hysteresis loop within the large one. This indicates the fatigue damage appearing during the stress cycle E to G.

The rainflow method counts the peaks and valleys of the stress profile such as the points indicated in Fig. 6.5. Hence the stress profile needs to be modified, leaving only the extreme points. The small stress cycles such as B to C in the figure is discarded since they have little effect on device lifetime. Considering the discussion in chapter 3, the die-attach solder is in the elastic zone when $\Delta T_j < 8$ ℃. Considering a safety factor of 50%, the temperature cycles that are less than 4 ℃ are considered in elastic zone and generate little damage to the solder, hence can be ignored.

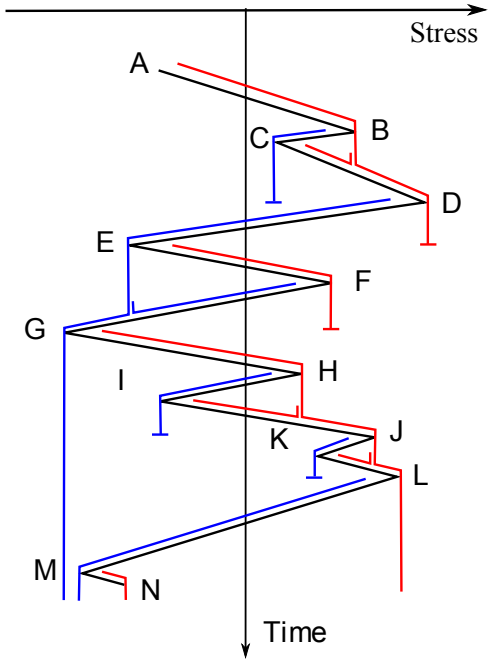


Figure 6.6: Original algorithm of rainflow cycle counting method, where red indicates the flow from valley and blue indicates the flow from peak

The original name of the rainflow method is called Pagoda Roof method. Fig. 6.6 shows the original algorithm of the rainflow method. The stress profile represents a series of roofs on which the water falls. The peak stress is on the right and valley is on the left. Each of the peak and valley are imagined to be the source of the water which drop down the pagoda. The algorithm is listed below:

- If the water drops from a peak of the stress profile:

  a. The drop will stop if it meets a peak larger than that of departure, for example, the blue drop starts at B and ends at D.

  b. It will stop if it meets the path traversed by another previously determined drop, like the blue drop from F terminates before G.

  c. The drop can fall on another roof and to continue to slip according to rules a and b.

- If the water drops from a valley of the stress profile:

  d. The fall will stop if it meets a valley deeper than that of departure, for example, the red drop starts from A and ends at E.

  e. The fall will stop if it crosses the path of a drop coming from a preceding valley, for example: the red drop starts from C terminates before D.

  f. The drop can fall on another roof and continue according to rules d and e.

The path way of the water fall is counted as a half cycle. For a sufficiently long stress profile, each peak generated half cycle will match a valley generated half cycle to form a whole cycle. The stress profile shown in Fig. 6.6 is counted as four whole cycles and five half cycles.

The original rainflow algorithm is convenient to carry out by hand which is only suitable for a short stress profile. A practical definition, which is suitable for computer analysis, is given by ASTM international Std [90]. The rules are given as following:

X denotes range under consideration; Y is the previous range adjacent to X; and S is the starting point in the history.

a. Read next peak or valley. If out of data, go to Step f.

b. If there are less than three points, go to Step a. Form ranges X and Y using the three most recent peaks and valleys that have not been discarded.

c. Compare the absolute values of ranges X and Y. If $X < Y$, go to step a; if $X \geq Y$, go to step d.

d. If Y contains the starting point S, go to step e; otherwise, count Y as one cycle; discard the peak and valley of Y; go to step b.

e. Count Y as one half cycle; discard the first point in range Y; move the starting point to the second point in range Y; go to step b.

f. Count each range that has not been previously counted as one half cycle.

Fig. 6.7 illustrates the counting process using practical rainflow definition. The details are explained in the following:

1. Start with three points, A, B, and C. The starting point is A, $X = |BC| < Y = |AB|$, read next point D.

2. Now $X = |CD| > Y = |BC|$ and Y do not contain the starting point A. Hence $|BC|$ is counted as one cycle, B and C are discarded.

3. Read point E since only two points left. Now $X = |DE| > Y = |AD|$ and Y contains the starting point. Hence $|AD|$ is counted as a half cycle, A is discarded and the starting point moves from A to D.

4. Read point F. $X = |EF| < Y = |DE|$, read next point G.

5. Now $X = |FG| > Y = |EF|$ and Y do not contain the starting point D. Hence $|EF|$ is counted as one cycle, E and F are discarded.

6. Read point H. $X = |GH| < Y = |DG|$, read next point I. $X = |HI| < Y = |GH|$, read next point J.

7. Now $X = |IJ| > Y = |HI|$ and Y do not contain the starting point D. Hence $|HI|$ is counted as one cycle, H and I are discarded.

8. Now $X = |GJ| < Y = |DG|$, read next point K. $X = |JK| < Y = |GJ|$, read next point L.

9. Now $X = |KL| > Y = |JK|$ and Y do not contain the starting point D. Hence $|JK|$ is counted as one cycle, J and K are discarded.

10. $X = |GL| < Y = |DG|$, read next point M. $X = |LM| < Y = |GL|$, read next point N.

11. $X = |MN| < Y = |LM|$, out of data, count $|DG|$, $|GL|$, $|LM|$, and $|MN|$ as half cycle and the counting process terminates.

It is obvious that the stress profile can be counted as four whole cycles and five half cycles, same as the results of the original algorithm. The Matlab codes, which can be found in Appendix C.3, are written to achieve this cycle counting method.

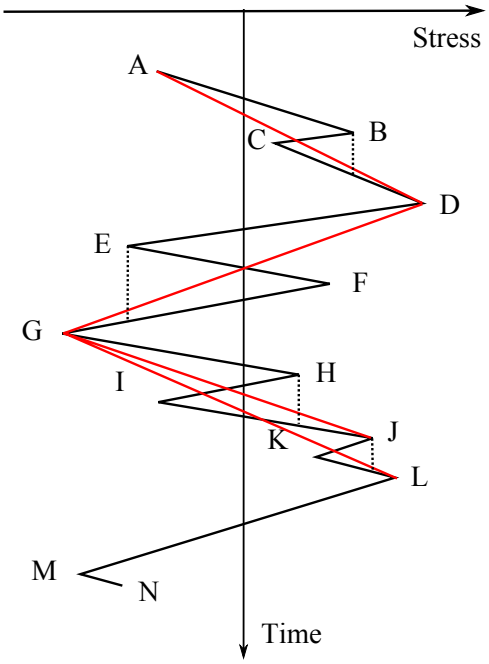


Figure 6.7: Illustration of practical rainflow algorithm.

6.3 Fatigue damage accumulation method

The first and most well known fatigue damage accumulation method was suggested by Palmgren in 1924 [91], this model assumes that the fatigue damage accumulates linearly. Miner expressed the linear damage concept in a mathematical equation as shown in Eqn. 6.6 [3], where r_i is the cycle ratio corresponding to the i^{th} load level, N_{fi} is the number of cycles to failure under the the i^{th} load level, n_i is the number of applied cycles in the i^{th} load level. This method is widely used by the reliability design engineers because of its simplicity. However, this method faces several shortcomings. The most important one is its load sequence independence, which is against the experimental evidence.

$$D = \sum r_i = \sum \frac{n_i}{N_{fi}} \tag{6.6}$$

Marco and Starkey proposed the first nonlinear load-dependent theory in 1954 [23], which

is represented by a exponential relationship based on their load sequence experiment results, as shown in Eqn. 6.7. Here, x_i is a variable quantity related to the i^{th} stress level. Fig. 6.8 shows the comparison of D-r (damage VS. cycle ratio) curves for the linear and exponential rule for damage accumulation. As reported by Marco and Starkey, x_i increases with the drop of the stress level. It is clear that for a low-to-high loading sequence such as the operation starting at S3 followed by S1 (follow the dotted line in Fig. 6.8), $\sum r_i = AC + BD > 1$; while for a completely reversed high-to-low loading sequence, $\sum r_i = AB + CD < 1$. This result agrees with the normal experimental evidence [61]. Furthermore, the nonlinear method also indicates that the damages caused by the same stress are different when the cycle ratio varies. The damage accumulates faster with a higher cycle ratio, which suits for the power cycling results shown in chapter 5. However, they did not explain how to calculate the stress related exponential factor x_i.

$$D_i = r_i^{x_i} \tag{6.7}$$

The power cycling data are modified and used to select a suitable accumulation method. Fig. 6.9 shows a typical D-r curve for both set of power cycling tests, where D=1 indicates a 50% increase of R_{th} according to the defined failure criteria. Clearly, this test results agree with Marco's results [23] and indicates that the nonlinear accumulation method fits better with the experimental data. Furthermore, x_i increases as ΔT_j decreases, which means the smaller stress cycle, the higher cycle ratio is needed to enter into the crack propagation process. Therefore, the nonlinear accumulation method is a better option for this particular work.

In order to use the nonlinear accumulation method, it is essential to find out how x_i varies with ΔT_j. The D-r curve of all test modules are fitted according to the nonlinear method shown in Eqn. 6.7. The fitting results are listed in Table. 6.2. Notice that the device lifetime increases with the drop of stress level according to the Coffin-Manson relationship, Eqn. 6.8 assumes x_i follows the similar exponential relationship. Here, c and d are constants

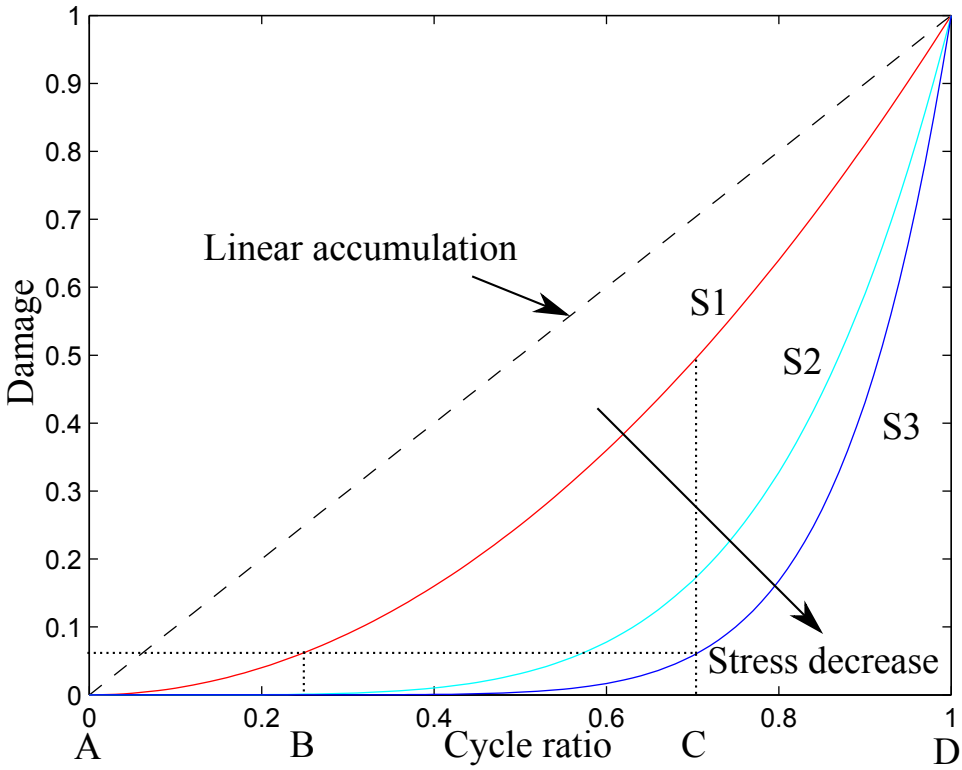


Figure 6.8: Comparison of D-r curves for the linear and nonlinear rule for damage accumulation. Where, S1, S2 and S3 indicates the nonlinear damage accumulation under different stress levels, while A, B, C and D indicates different lifetime point.

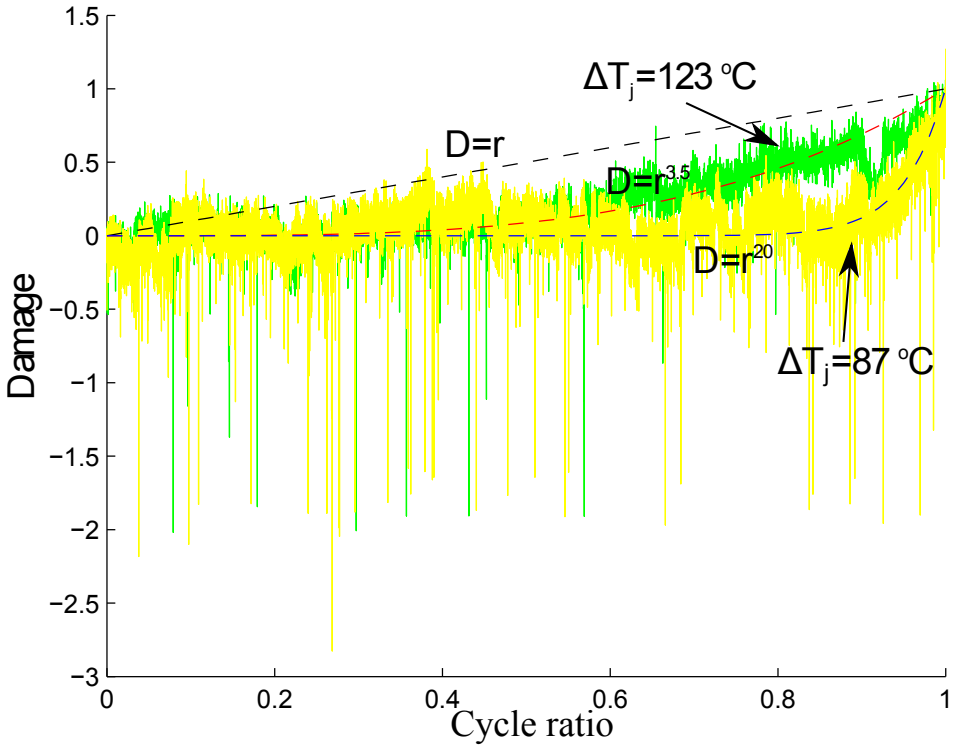


Figure 6.9: Typical D-r curves of both sets of power cycling tests.

which need to be extracted from the results shown in Table. 6.2. According to the extracted results, the average x_i is 12.3 for test 1 and 2.85 for test 2. Substituting these values into Eqn. 6.8, the two constants can be calculated as, $c = 1.7636 \times 10^9$ and $d = -4.2067$. The exponential term d is very close to the Coffin-Manson coefficient b' shown in Table 6.1, which supports the assumption made above. This phenomenon can be explained by the crack growth theory [92] since cracks are directly related to damage. This theory divides the fatigue damage into two phases, crack initiation and crack propagation. During the crack initiation phase, the crack does not exist or it is too tiny to cause damage. Therefore, the damage accumulation in this phase are nearly constant and can be thought of as being equal to zero. After a certain number of stress cycles, the crack develops to a critical length and starts to propagate, which represents the damage accumulation. In the case of solder fatigue, the increase of thermal resistance indicates the crack growth and damage accumulation. A higher stress cycle amplitude leads to a faster crack initiation and damage accumulation, as the test results shown in Fig. 6.9.

$$x_i = c\Delta T_j^d \tag{6.8}$$

In order to apply the accumulation method in a real mission profile which contains multi-level stresses, the assumption of linear damage accumulation in one cycle is made. From Fig. 6.8, it is clear that for a constant stress level, the damage caused by each cycle depends on the its cycle ratio. Therefore, the accumulated damage in each cycle can be calculated by the steps listed below.

1. Calculate the cycle ratio (r_i) from the accumulated damage of the previous cycles (D_{i-1}) by $r_i = D_{i-1}^{1/x_i} + \Delta r_i$, where $\Delta r_i = n_i/N_{fi}$. This is because r_i varies as the stress changes. For example, when the stress changes from S1 to S3 at point B as shown in Fig. 6.8, r_i moves from point B to point C and keeps the accumulated damage the same.

Table 6.2: Extracted x_i for two sets of power cycling test.

IGBT NO	x_i	**IGBT NO**	x_i
1.1	2.698	2.1	3.362
1.2	21.82	2.2	2.836
1.3	4.716	2.3	2.478
1.4	6.738	2.4	5.523
1.5	6.198	2.5	3.764
1.6	4.166	2.6	2.410
1.7	4.043	2.7	2.832
1.8	19.65	2.8	2.319
1.9	20.28	2.9	2.617
1.10	24.02	2.10	2.521
1.11	N/A	2.11	2.445
1.12	13.64	2.12	3.217
1.13	8.761	2.13	2.552
1.14	20.13	2.14	2.552
1.15	24.01	2.15	2.579
1.16	6.028	2.16	2.845
1.17	3.544	2.17	2.705
1.18	N/A	2.18	2.213
1.19	7.868	2.19	2.453
1.20	21.82	2.20	1.831

2. Calculate the damage due to the i^{th} cycle by $\Delta D_i = x_i \cdot r_i^{x_i - 1} \cdot \Delta r_i$ according to the linear assumption.

3. Accumulate the damage according to $D_i = D_{i-1} + \Delta D_i$.

However, the above method faces a practical problem. When ΔT_j is relatively small, both x_i and N_{fi} are huge which may lead to a tiny ΔD_i that is below the minimum calculation range of Matlab (1×10^{-323}). This problem is extremely serious for a small cycle ratio, and will lead to a zero damage accumulation until a large temperature cycle occur. For example the first cycle of the mission profile, $\Delta T_j = 40$ ℃ and $r_1 = \Delta r_1 = 1/N_{f1}$. According to Eqns. 6.7 and 6.8, it can be calculated that $x_1 = 321.38$, $N_{f1} = 1.7937 \times 10^6$, and consequently $D_1 = 9.1 \times 10^{-2008}$, which is treated as a zero in Matlab.

Fig. 6.10 shows the D-r curves of $\Delta T_j = 40$ ℃ and $\Delta T_j = 20$ ℃. It is clear that the accumulated damage of both curves are tiny when the accumulated damage is small. Therefore, it is appropriate to assume that no damage had accumulated until it has reached a threshold value. In this region, r accumulates linearly without causing damage. The threshold is defined as $D = 1 \times 10^{-300}$ according to the minimum calculation range of Matlab. For a given mission profile, the accumulated cycle ratio per mission profile can be calculated. The D-r curve of the maximum ΔT_j in this mission profile is used to locate the r where damage accumulates to the threshold. The damage accumulation process starts from the threshold and is simulated according to the steps listed above. The Matlab code for the nonlinear damage accumulation is available in Appendix C.4.

In order to check the validity of this accumulation method, two sets of mission profile have been designed for the simulation. The first one is constant stress profile which $\Delta T_j = 80$ ℃. In this test, the N_f calculated by Coffin-Manson law is 82312 cycles while the N_f simulated by the nonlinear accumulation code is 82398 cycles. The difference is due to the linear damage accumulation assumption within one cycle. However, the error is only 0.1% which is negligible. Fig. 6.11 shows the characteristics of the nonlinear damage accumulation method under the constant stress test. It is clear that Δr is constant and hence r accumulates linearly,

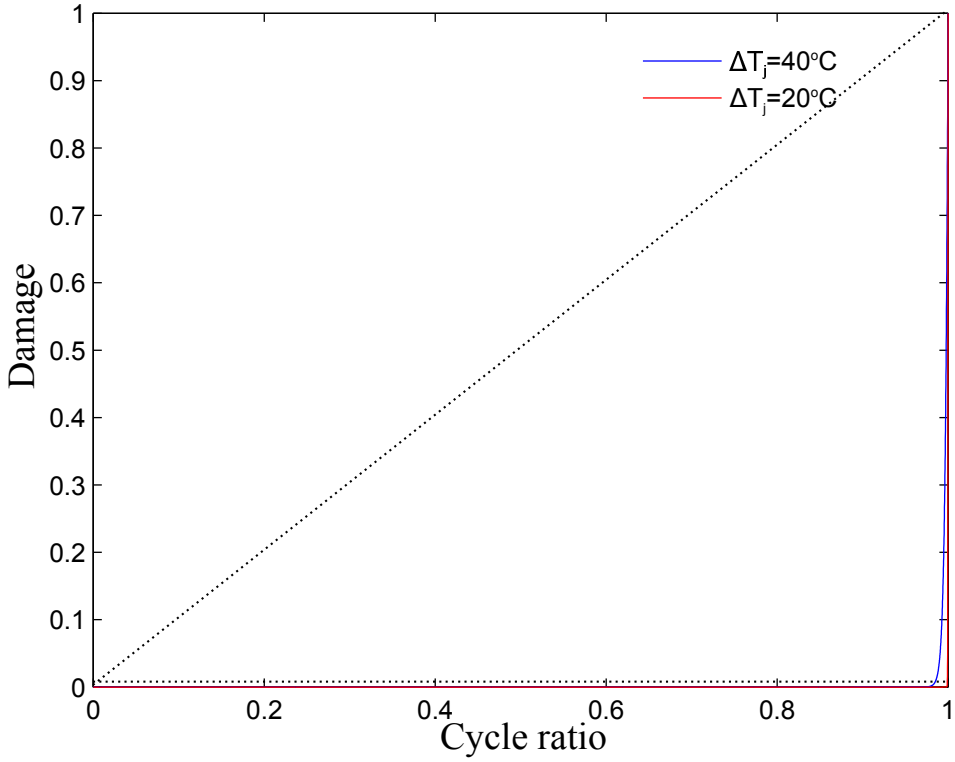


Figure 6.10: D-r curves for small temperature cycles.

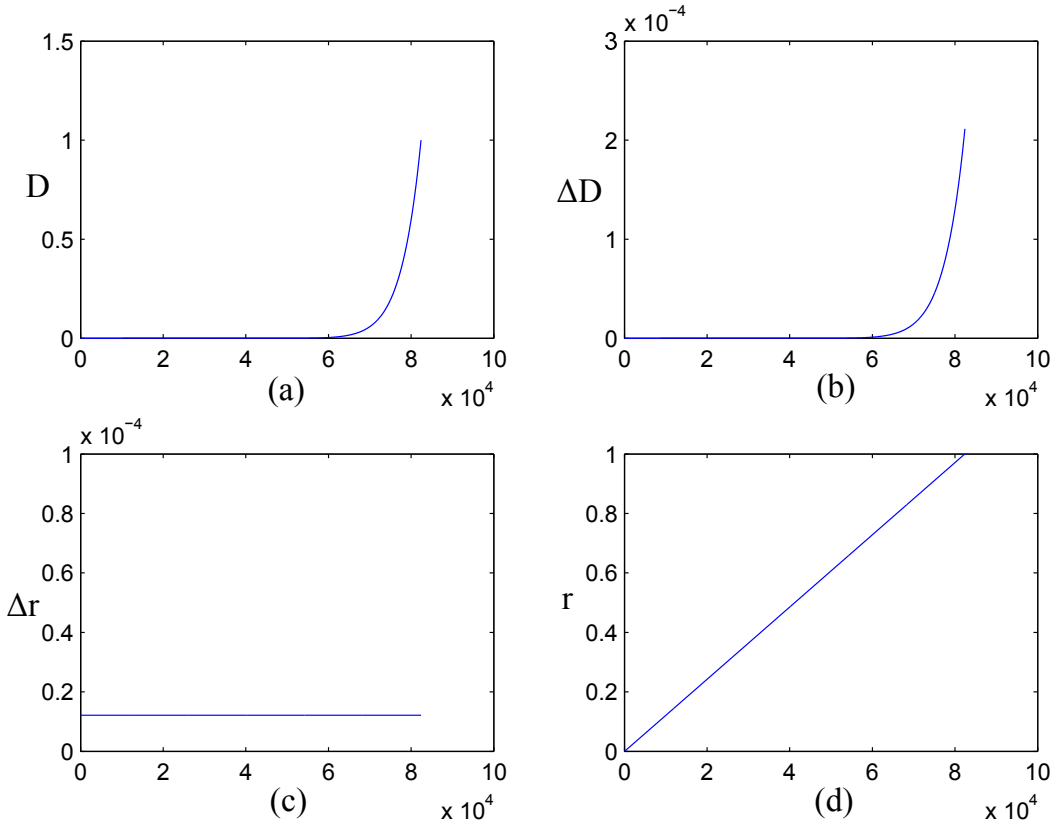


Figure 6.11: Validation of nonlinear accumulation method with constant stress, (a) accumulated damage during lifetime, (b) damage per cycle during lifetime, (c) Δr per cycle during lifetime, (d) accumulated r during lifetime. x axis indicates the number of cycles.

since the stress cycle is constant. The damage due to each cycle increases as r increase, which agrees with the experimental observations as shown in Fig. 5.17. The accumulated damage follows the nonlinear rule as shown in Eqn. 6.7 with x = 17.4. These simulation results show that the accumulation method and assumptions are valid for the constant stress cycle load.

The second profile is a two-step stress cycle which consists of a big stress cycle $\Delta T_j =$ 80 ℃ followed by a small stress cycle $\Delta T_j = 60$ ℃. This profile represents the simplest multi-level stress load. According to Eqn. 6.5, N_f is 8.23×10^4 and 2.9573×10^5 for $\Delta T_j = 80$ ℃ and 60 ℃ respectively. Hence the N_f calculated by the linear damage accumulation method is 128780 cycles while the N_f simulated by the nonlinear accumulation code is 109491 cycles.

It is clear that N_f is dominated by the bigger stress cycle. Fig. 6.12 shows the characteristics of the nonlinear damage accumulation method under the two-step load test. The D-r curve of the combined stress is located between the D-r curves of big and small load cycles and is closer to the D-r curve of the big load cycle. The damage per cycle is calculated by $\Delta D_i = x_i \cdot r_i^{x_i-1} \cdot \Delta r_i$. When r_i is small, i.e. in crack initiation region, the damage caused by a large stress is much greater than that caused by a small stress. This is because x_i is 17.4 and 58.4 for large and small stresses respectively. For small r_i, the term $r_i^{x_i-1}$ dominates the scale of ΔD. As r_i approaching 1, the term $r_i^{x_i-1}$ approximates to one, hence ΔD is dominated by $x_i \cdot \Delta r_i$. Although Δr_i for the small stress is less than that for the big stress, its x_i is greater. For this test, $\Delta D_{80}/\Delta D_{60} = 173$ at the beginning and $\Delta D_{80}/\Delta D_{60} = 1.07$ at the end. Therefore, ΔD caused by different stresses are similar in the crack propagation region, as shown in Fig. 6.12(b). The lifetime is dominated by bigger stress since most of the lifetime is in the crack initiation region. As the damage accumulates, r for the two stresses follows different curves as shown in Fig. 6.12(d). This is because r for the smaller stress is greater than the r for bigger stress when D is the same, which is clearly demonstrated in Fig. 6.8. The test results show the desired characteristics and support the applicability of the nonlinear damage accumulation method in the multi-level stress load.

As damage accumulates, the thermal resistance increases which could lead to the increase of the junction temperature. Therefore for an actual inverter, the temperature profile is not the same even the electrical profile stays unchanged. The increase of temperature not only leads to the increase of present cyclic stresses but also pushed some temperature cycles from the elastic zone into the plastic zone, which leads to the rise in the number of cycles in one mission profile. This effect will accelerate the crack propagation and reduce the inverter lifetime.

In order to represent this effect, a feedback loop is proposed in this work to update the temperature profile, as shown in the lifetime damage model Fig. 1.7. However, it is not necessary to update the temperature in every mission profile since nearly no damage

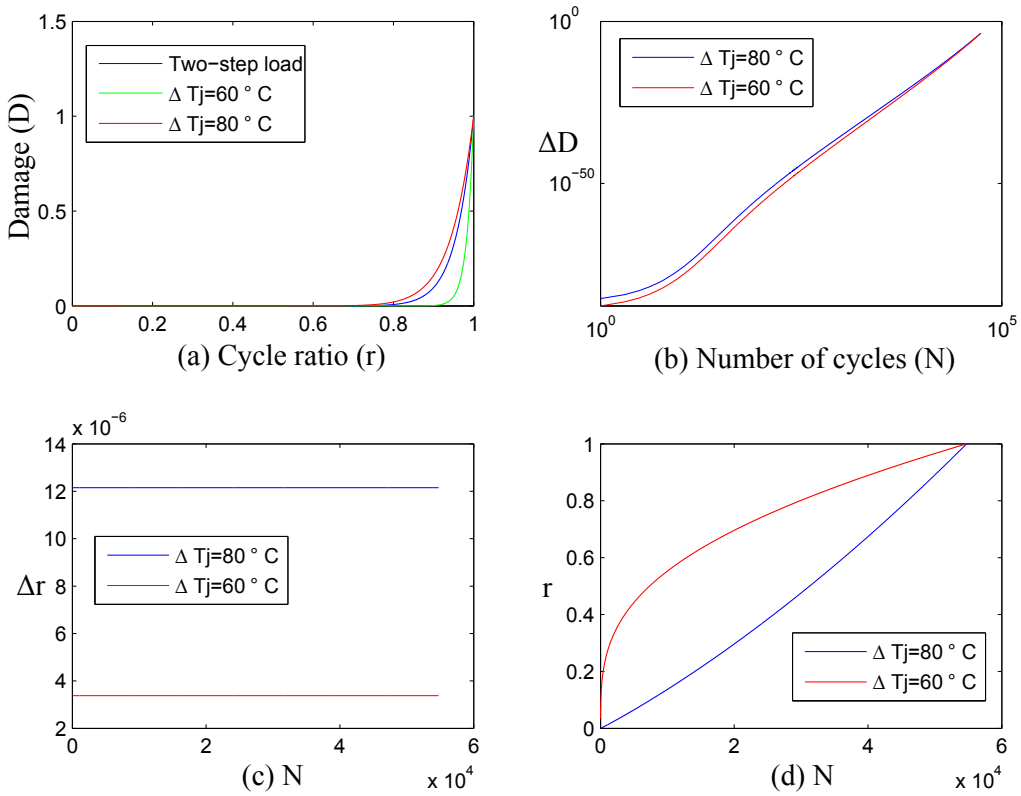


Figure 6.12: Validation of nonlinear accumulation method with two-step stress, (a) D-r curves for the nonlinear accumulated two-step load, $\Delta T_j = 80$ ℃, and $\Delta T_j = 60$ ℃, (b) damage per cycle during lifetime, (c) Δr per cycle during lifetime, (d) accumulated r during lifetime.

has accumulated in the crack initiation region and the electrothermal simulation is time consuming. Therefore, an alternative is to find out a reasonable percentage of damage for the feedback loop. Assuming the thermal capacitance, T_c and the device power dissipation are constant, the maximum possible T_j increase is the same percentage as the increase of R_{th} at stable state, according to Eqn. 4.6. Obviously the percentage increase of T_j is smaller than that of R_{th} if the system is a in transient state. Therefore a 10% increase of R_{th} (which can lead to a 10% maximum possible T_j increase) is chosen as a reasonable step to update the temperature profile in this work. The failure ($D = 1$) is defined as a 50% increase of R_{th}. Hence it is only necessary to run the electrothermal model five times at R_{th}, $1.1R_{th}$, $1.2R_{th}$, $1.3R_{th}$ and $1.4R_{th}$ respectively, which is time efficient. However, in the controlled ΔT_j tests, this feedback affect do not appear. The comparison of these three different accumulation method will be presented in the next chapter.

Notice that the thermal capacitance is assumed to keep constant in this work, which might affect the accuracy of the simulation. In the simulation of a real system, a more accurate thermal model, such as the one developed by Ian Swan [93], can always be used to obtain more accurate results.

6.4 Conclusions

This chapter developed a die-attach solder fatigue damage model based on the Coffin-Manson law and the statistical analysis of the power cycling data. Different cycle counting methods are discussed and the rainflow method was selected. Two damage accumulation methods have been discussed and compared. Most power module lifetime prediction models available in the literature use Miner's rule to accumulate damage linearly. This method does not correlate well when compared with the experimental results. A more accurate nonlinear accumulation method is promoted and its multi-level stress test shows a more conservative lifetime estimation than the linear accumulation method.

All the blocks in the lifetime estimation model have been demonstrated. Chapter 2 introduced the basic concept of the device switching model. Chapter 4 explained the configuration of the electrothermal model. This chapter derived the fatigue damage model from the power cycling results listed in Chapter 5, introduced the cycle counting method and damage accumulation method. Having presented each of these techniques, the inverter lifetime prediction model can be implemented. A case study of inverter lifetime estimation for wind energy application is designed in the next chapter to show an example of the implementation of the inverter lifetime damage model.

Chapter

7 Case Study: Wind Power Application

Previous chapters explained the techniques needed for the implementation of the inverter lifetime damage model. These techniques are used in this chapter to demonstrate the application of the model in a wind power system. After a careful configuration of this model, the lifetime of the inverter can be estimated with a given wind speed profile. Due to the lack of real system data, the lifetime model build in this chapter uses the same electrothermal model and fatigue damage model developed from previous chapters. The same technique can be used to simulate a real wind turbine system if required data is available.

7.1 Inverter Electrothermal Modelling for Wind Power Application

The electrical energy generated from renewable sources cannot be sent to the grid until it has been modified to the required conditions. Power converters are widely installed to accomplish this task. There are several special operating conditions for the inverter used in the wind power application. First of all, the voltage of DC bus which is involved in most wind power converters is kept almost constant during normal operation [94]. Hence the input

of the inverter can be considered as a constant DC voltage source. Secondly, its output is connected to the grid which can be considered as a three-phase voltage source. This means the load voltage is constant and the input power fluctuation due to the variation of wind speed can only be represented by the inverter output current. Therefore, a current control technique is needed to regulate the inverter output current according to the power generated by the wind turbine.

7.1.1 Wind Power Generation System

There are many topologies available for the wind power generation system. They can be sorted into three groups, the systems without power electronics, the systems with partially rated power electronics, and the systems with fully rated power electronics. Fig. 7.1(a) shows the system without power electronics, which normally uses induction generator to work at a fixed rotation speed since its output frequency cannot be modified and need to be the same as the grid frequency. The input power of this system is limited by the wind turbine pitch control [95], therefore this type of system cannot optimise the utility of wind energy. Furthermore, any fluctuation in wind speed naturally causes the mechanical power of the turbine to vary. This could lead to the torque fluctuation of the turbine rotor and stress the mechanical component such as gear box.

Fig. 7.1(b) shows a widely used system topology, doubly-fed induction generator (DFIG), in the second category. Partially rated power converters are employed hence greatly improved the system control performance. By using the DFIG, the system is able to operate at variable rotor speed while the amplitude and frequency of the generated voltage remain constant. This is achieved by adjusting the amplitude and frequency of the AC currents that fed into the generator rotor windings. With the ability of variable speed operation, the torque fluctuation at the mechanical component is reduced and more power can be generated from wind energy. Furthermore, the power rating of the converter is about 30% of the nominal generator output power, which saves the cost for power electronics and reduces the power

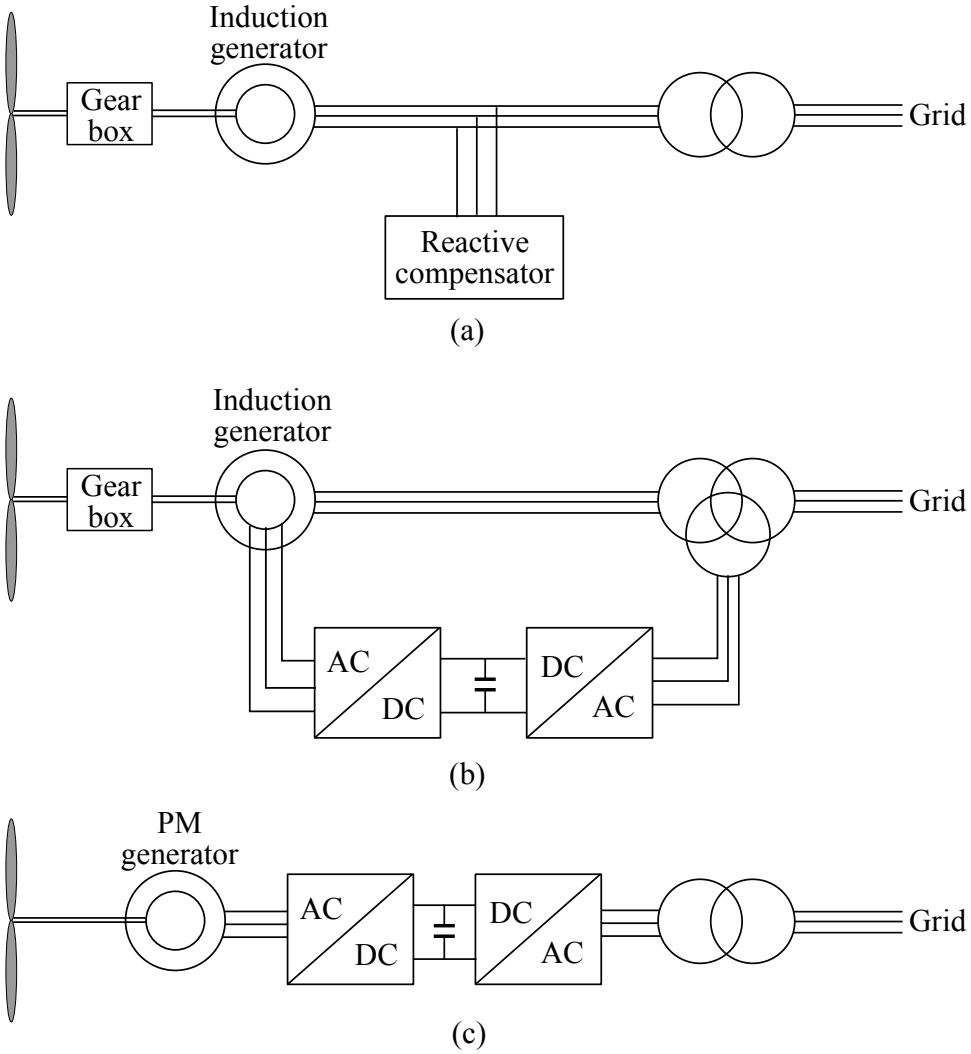


Figure 7.1: Different topologies of wind power generation system. (a) topology without power electronics, (b) topology with partially rated power electronics, (c) topology with fully rated power electronics.

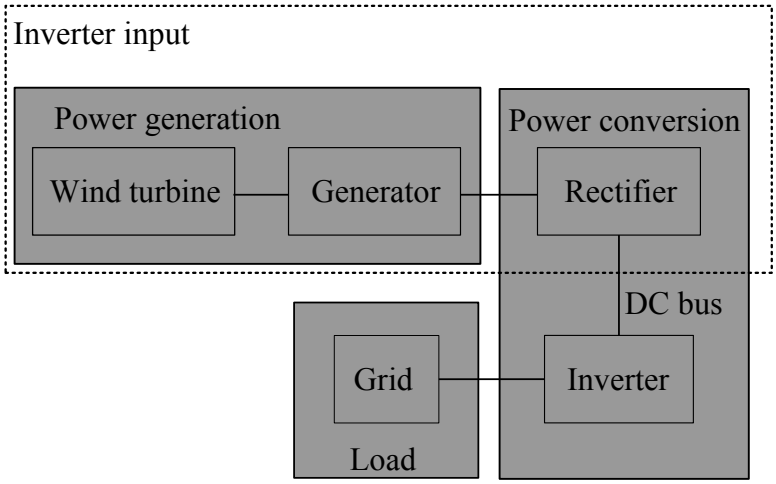


Figure 7.2: Basic structure of wind power generation system.

loss during power conversion. However, the DFIG requires a complex power conversion and control circuit and the slip rings used for the rotor excitation requires periodic maintenance, which is not desirable for the off-shore operation.

Permanent magnet (PM) generator can be used to avoid the excitation circuit as well as the slip rings. The topology of this system is shown in Fig. 7.1(c), which is sorted into the third category. This system is able to operate at variable speed as well, but the amplitude and frequency of the generator output voltage vary as the wind speed fluctuates. Hence its power conversion unit needs to be the same rating as the generator. Furthermore, the increasing cost of PM elements also limits the application of the PM generator. However, this type of system eliminates the gear box and slip rings which have high failure rates, hence improve its reliability.

This work focuses on the reliability prediction of power converters thus the PM generator topology is selected for the simulation due to its simplicity in control and electrical circuit. The basic structure of such wind power generation system consists of the power generation unit, the power conversion unit, and the load, as shown in Fig. 7.2. As the DC bus voltage is kept almost constant, the output of the rectifier can be considered as a constant voltage source, whose output current is determined by the output power of the PM

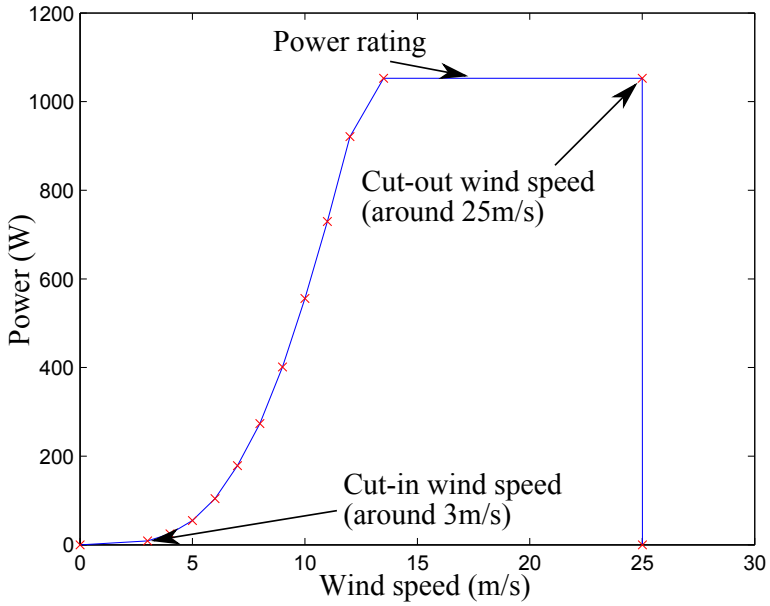


Figure 7.3: Inverter input power curve.

generator. Therefore it is essential to obtain the output power variation due to the wind speed fluctuation.

The output power of the PM generator depends on the power generated by the wind turbine and the efficiency of the generator. The wind turbine output power is determined by the wind speed and can be obtained from its power curve. The power curve of an actual wind turbine (RES 1 MW) is scaled down to the power rating of the test modules (SKM50GB123D), as shown in Fig. 7.3. The modified power curve is used as a LUT to generate the wind turbine output power according to wind speed profile. The efficiency of the power generation and conversion unit is assumed to be 90%. Since the DC bus voltage is constant, the wind turbine output power can be used to derive the inverter output current. Therefore the detailed modelling of the wind turbine, the generator, and the rectifier can be neglected.

The main electrical circuit of the grid connected voltage source inverter model is designed and shown in Fig. 7.4. This inverter model is similar to the general purpose inverter model

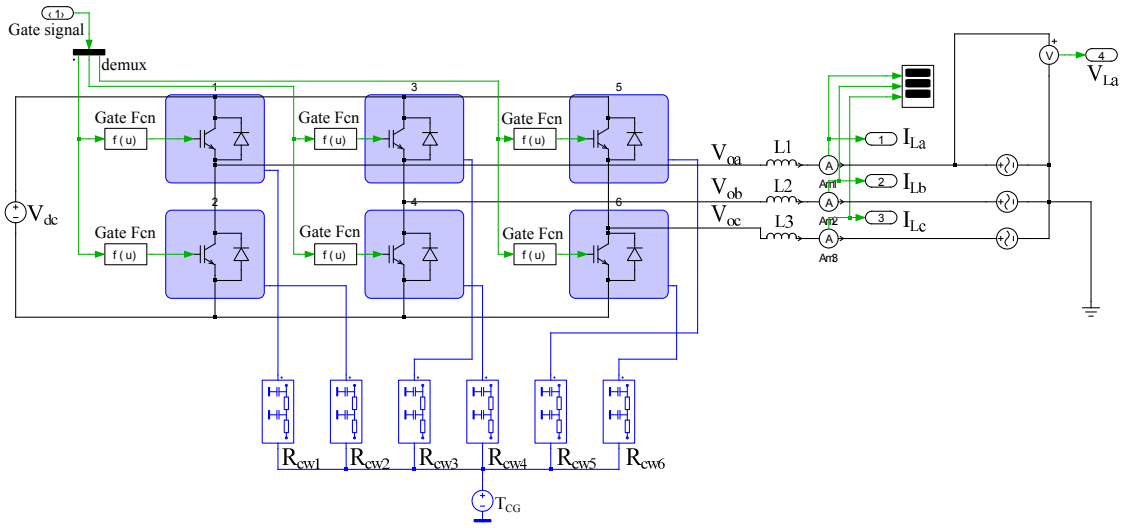


Figure 7.4: PLECS model of the grid connected voltage source inverter.

introduced in chapter 4, with some small modifications on the circuit. The resistive load is changed to a constant three phase AC voltage source which represents the ideal grid. The DC bus voltage is set to be 800 V and the three phase grid voltage is set to be 220 V. The capacitors are removed since they tend to short circuit the AC voltage source. Inductors (0.01 H) are used to ensure the current conducts continuously. They also work as filters to reduce the harmonics generated by the inverter. When applying this model to a real application, the ambient temperature may change, this can be modified by change the constant temperature sources T_{CG} in Fig. 7.4 to an variable temperature source which value is the recorded ambient temperature.

7.1.2 Inverter Current Control Technique

Comparing the open loop voltage source PWM inverters, the inverter with close loop current control offers the ability to control the instantaneous current waveform with high accuracy, which gives the advantage of lower distortion and harmonic noise. There are three major techniques used for controlling the output current of a voltage source inverter. They are hysteresis control [96], ramp comparison [97] and predictive control [98].

The hysteresis control is the simplest method among these three. Its principle is to switch the inverter leg in the negative/positive direction when the corresponding current is greater/less than the reference current by the hysteresis band. Therefore the hysteresis band specifies the maximum current ripple. By applying this control method, the inverter switching frequency varies over a fundamental inverter period which is not desirable for eliminating the harmonics.

In the case of ramp comparison control, the current error is compared to a triangle waveform. The inverter leg is switched in the positive/negative direction if the current error is greater/less than the triangle waveform. Compared with the hysteresis controller, the ramp controller change the constant hysteresis band to a fixed frequency triangle waveform and can be thought of producing an sine-triangle PWM with the current error to be the modulating function. Therefore the inverter switches at the frequency of the triangle waveform and produces well defined harmonics. However, there is an inherent magnitude and phase error between the reference current and line current.

The predictive current controller calculates the inverter output voltages that required to force the measured load current to follow the reference current. Hence it needs more calculations and requires a good knowledge of the system parameters. This controller offers a constant inverter switching frequency and the best potential for precise current control [99], hence it is selected in this work to control the inverter operation.

The key of the predictive current control strategy is to use the measured results of previous switching cycle to estimate both the next inverter output voltage required to support the reference current. According to Fig. 7.4, the voltage equation of this circuit can be obtained as shown in Eqn. 7.1. Here, V_O is the inverter output voltage, V_L and I_L are the load voltage and current respectively as defined in this figure. Assuming the inverter operates with a constant switching period (T_S), Eqn. 7.1 can be written in a discrete form during the switching period [n, n+1] as shown in Eqn. 7.2. Where $V_{O\_ave}[n]$ and $V_{L\_ave}[n]$ are the average inverter output voltage and grid voltage over the switching period [n, n+1]

respectively, $I_L[n+1]$ and $I_L[n]$ are the instantaneous current at sampling point [n+1] and [n] respectively. The aim of the controller is to force the load current at point [n+1] to be equal to the reference current during the switching period [n, n+1]. Therefore the required predictive average inverter output voltage is defined as shown in Eqn. 7.3, where $I_{ref}[n+1]$ is the instantaneous reference current at sampling point [n+1].

$$V_O = V_L + L\frac{dI_L}{dt} \tag{7.1}$$

$$V_{O\_ave}[n] = V_{L\_ave}[n] + L\frac{I_L[n+1] - I_L[n]}{T_S} \tag{7.2}$$

$$V_{O\_ave}[n] = V_{L\_ave}[n] + L\frac{I_{ref}[n+1] - I_L[n]}{T_S} \tag{7.3}$$

In Eqn. 7.3, $V_{L\_ave}[n]$ and $I_L[n]$ are the unknown parameters that need to be calculated from previous measurement. In order to predict the average grid voltage ($V_{L\_ave}[n]$), the change of the grid voltage during [n-2, n+1] is assumed to be linear. $V_{L\_ave}[n]$ can then be estimated from previously measured voltages at sampling point [n-2] and [n-1] using simple linear extrapolation, as illustrated in Fig. 7.5. The extrapolation result is shown in Eqn. 7.4. While $I_L[n]$ can be estimated by adding the predicted current change during the switching period [n-1, n] to the measured current at sampling point [n-1], as shown in Eqn. 7.5. Substituting Eqns. 7.4 and 7.5 into 7.3, the required average inverter output voltage during period [n, n+1] can be expressed as shown in Eqn. 7.6.

$$V_{L\_ave}[n] = 2.5V_L[n-1] - 1.5V_L[n-2] \tag{7.4}$$

$$I_L[n] = I_L[n-1] + \frac{T_S}{L}\left(V_{O\_ave}[n-1] - \frac{3V_L[n-1] - V_L[n-2]}{2}\right) \tag{7.5}$$

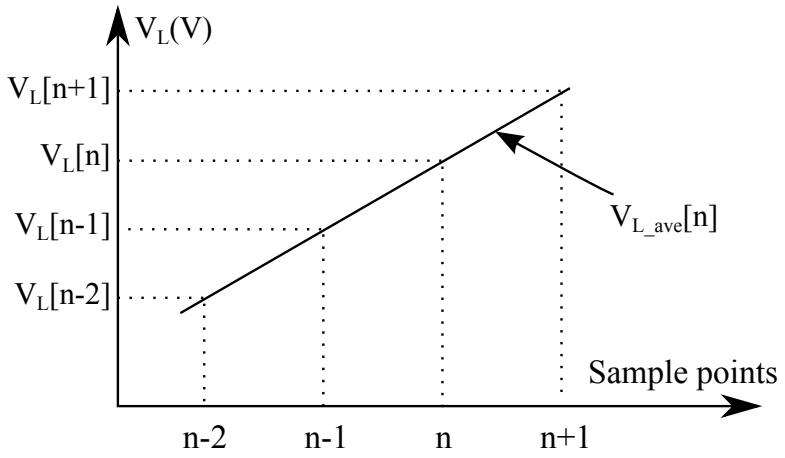


Figure 7.5: Linear assumption of predictive current control.

$$V_{O\_ave}[n] = 4V_L[n-1] - 2V_L[n-2] - V_{O\_ave}[n-1] + L\frac{I_{ref}[n+1] - I_L[n-1]}{T_S} \tag{7.6}$$

The control block of predictive controller is built by implementing Eqn. 7.6 in Matlab Simulink, as shown in Fig. 7.6. There are four inputs for this controller. The inverter three-phase output current (I_{3ph}) and inverter phase a output voltage (V_a) are measured from the electrical simulation circuit as shown in Fig. 7.4. Angle refers to the grid phase angle (θ), which is calculated from the simulation time. Assuming the phase angle is zero when the simulation started, the phase angle can be calculated by the Matlab command as shown in Eqn. 7.7, where f_s is the grid frequency and t_s is the simulation time. The amplitude of I_{ref} is estimated from the modified turbine power curve LUT, wind speed profile, and the grid voltage. As shown in Fig. 7.6, the previous value of a parameter is calculated by phase shifting. For example, $V_L[n-1]$ and $V_L[n-2]$ is calculated by shifting the current value of V_L back by a sampling period T_S and $2T_S$ respectively. The predictive current $I_{ref}[n+1]$ is calculated by shifting its current value forward by T_S. The desirable inverter output voltage is calculated by applying Eqn. 7.6. The result is then transformed into the stationary frame and is used as the input of the SVPWM generator. The predictive current control strategy

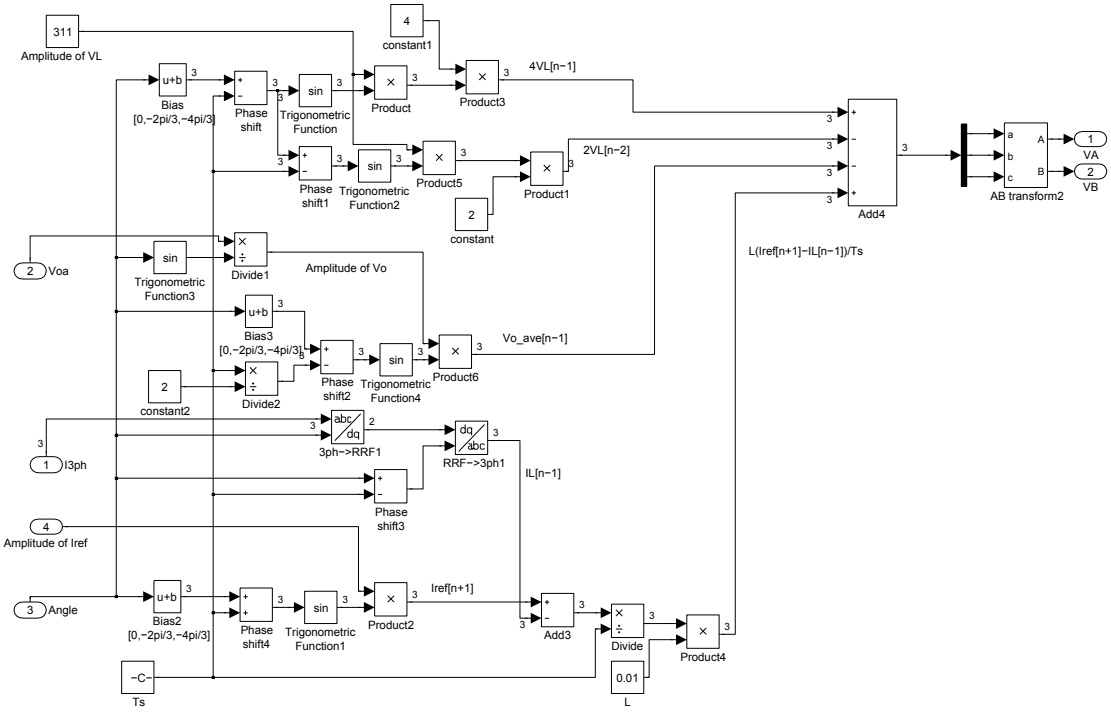


Figure 7.6: Block diagram of the predictive controller build in Matlab Simulink.

will introduce a constant current phase delay (T_S) [99]. However it is ignorable when the switching frequency is much higher than the fundamental frequency.

$$\theta = mod(2\pi f_s t_s, 2\pi) \tag{7.7}$$

7.1.3 Inverter Electrothermal Simulation

A modified inverter electrothermal model, which is based on the simulation technique introduced in chapter 4, is used to simulate the IGBT junction temperature. Both the power loss LUTs and the thermal models are kept the same as that in chapter 4. However, they can be modified to fit a real inverter system for different applications. A wind speed profile which is based on a real wind speed data is used as the input of this system, as shown in Fig. 7.7(a). By applying the current control technique which reference current is estimated

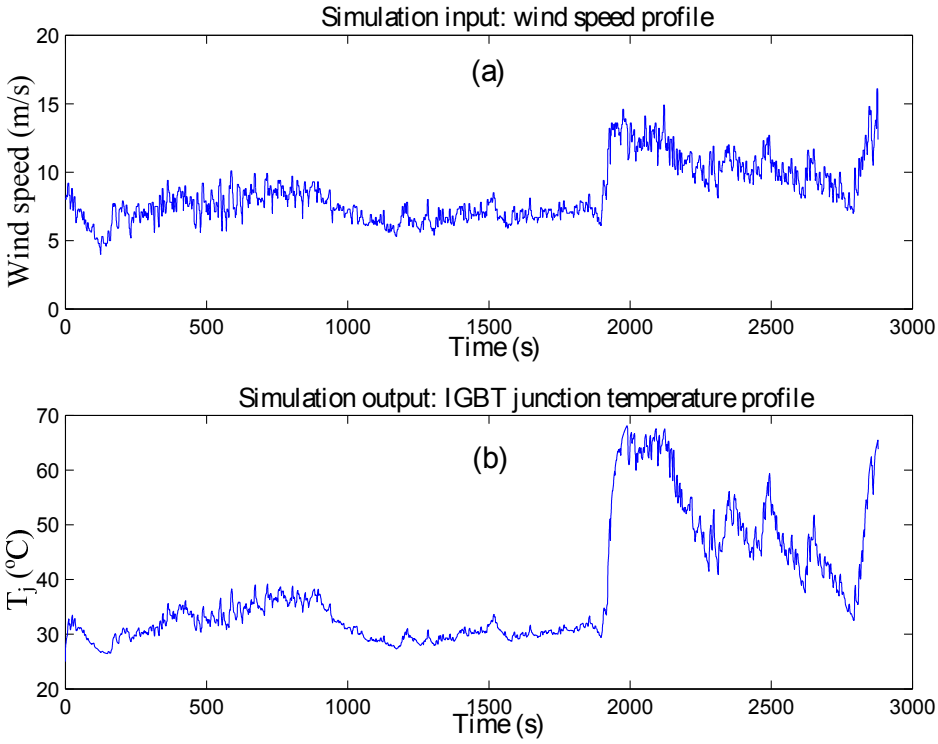


Figure 7.7: Simulation (a) input and (b) output.

from the wind speed, the IGBT junction temperature can be simulated and shown in Fig. 7.7(b). The simulation result shows that the IGBT junction temperature increases/decreases as the wind speed increases/decreases. This is because more power is generated by the wind turbine when the wind speed increases, which lead to a higher inverter current thus more heat is dissipated by the IGBT. Furthermore, the wind fluctuations at higher speed (after 2000 s) could lead to greater temperature variations. This is because the power generated from wind energy increases with the wind speed by a power relationship as shown in Fig. 7.3. The wind fluctuation at higher speed will cause a higher power fluctuation which could lead to a greater IGBT junction temperature variation.

7.2 Inverter Lifetime Prediction

The lifetime prediction of the voltage source inverter is based on the fatigue damage model obtained from the power cycling test and the IGBT junction temperature profile generated from the inverter electrothermal simulation. The temperature profile is modified, leaving only the extreme points and then counted by the rainflow method. The damage caused by each cycle is calculated by the fatigue damage model and is accumulated by linear and nonlinear methods separately. The estimated result is compared and discussed. The necessity of the feedback loop is discussed for both linear and nonlinear accumulation method.

7.2.1 Rainflow Cycle Counting of the Temperature Profile

The rainflow method is applied to count the number of cycles within one temperature profile. The temperature profile is modified before the cycle counting process since only the extreme points are needed. A matlab code is designed for this purpose and is presented in Appendix C.1. The temperature profile before and after this modification is shown in Fig. 7.8. From this figure it is clear that most characteristics of the original temperature profile can be represented by the modified profile, including the temperature cycle and the time stamps of the extreme points. The temperature cycles smaller than 1 ℃ are filtered in this process since they are too small to cause plastic deformation.

Rainflow method is applied to the modified temperature profile to count the valid number of cycles. As discussed in chapter 6, the temperature cycles less then 4 ℃ are discarded in this process since they are in elastic deformation region. The counted cycles are sorted in a histogram diagram according to their ΔT_j, as shown in Fig. 7.9(a). The number of cycles to failure for each cycle is calculated and are accumulated based on the assumption that damage accumulates linearly within one mission profile. The rainflow damage histogram is then plotted as shown in Fig. 7.9(b). From this figure, it is clear that although the majority temperature cycles are small, the damage caused by them only account for a very small

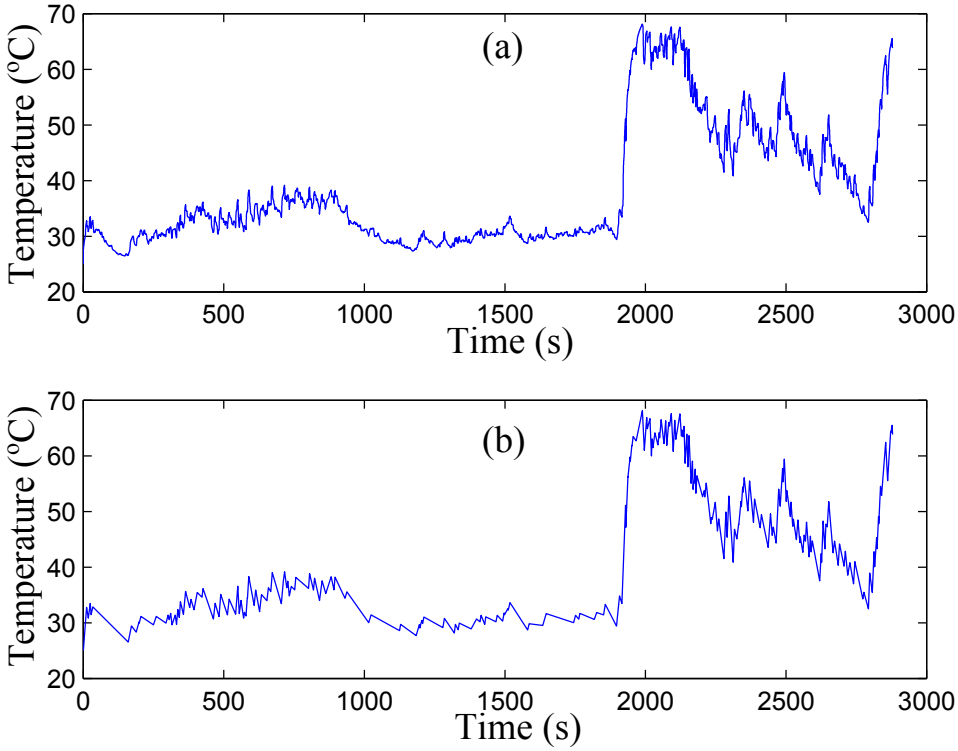


Figure 7.8: (a) Original IGBT junction temperature profile and (b) extreme points of the temperature profile.

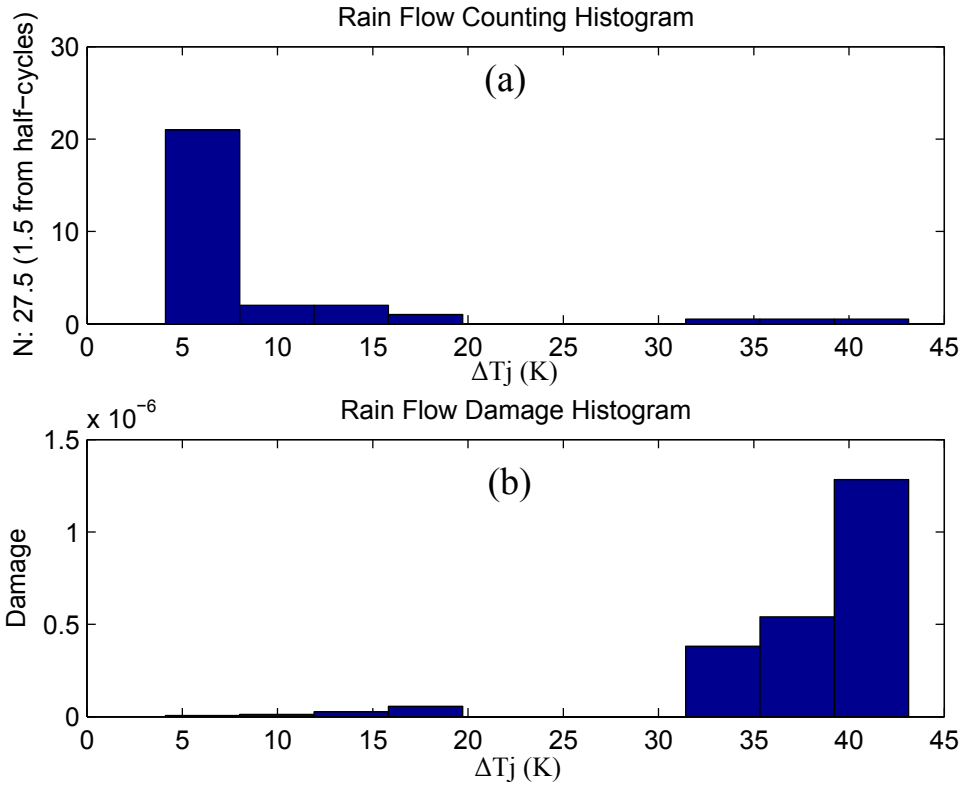


Figure 7.9: (a) Rainflow cycle counting results for one temperature profile and (b) the damage caused by the temperature cycles within one profile.

portion. Most of the damage is caused by the temperature cycles with large ΔT_j. This means that the inverter lifetime is dominated by the effects of the large temperature cycles.

7.2.2 Damage Accumulation

The damage accumulation method greatly affects the accuracy of the lifetime prediction. Two methods are discussed in chapter 6. From the experimental results, it is clear that the nonlinear method provides a better fit. Furthermore, according to the two-level stress simulation, the nonlinear method gives a more conservative prediction. The temperature profile shown in Fig. 7.7(b) consists of multi-level temperature cycles. The predicted lifetime by the the two damage accumulation methods are compared and discussed in this section.

The lifetime estimation according to Miner's law is quite straightforward. Since the damage accumulates linearly, the damage caused by each temperature profile is the same. Hence the lifetime can be calculated from the number of mission profile it can survive. Adding all the damage components in the rainflow damage histogram, the damage caused by one temperature profile can be obtained as 2.3084×10^{-6}. Therefore, the device lifetime can be estimated by Eqn. 7.8, where t_m is the length of the temperature profile and D_m is the damage caused by one profile.

$$Lifetime = \frac{t_m}{D_m} = \frac{2879\ seconds}{2.3084 \times 10^{-6}} = 1.2472 \times 10^9\ seconds = 39.5\ years \tag{7.8}$$

Applying the nonlinear accumulation method is a bit more complicated than the linear method. The damage accumulation process is divided into two phases, crack initiation and crack propagation. In the crack initiation phase, the damage is assumed to accumulate linearly. While in the crack propagation phase, the nonlinear method is applied. The threshold damage is 1×10^{-300}, which is the minimum value that Matlab is able to calculate. The counted cycles is analysed by the nonlinear accumulation code. The result shows that it takes 74111 profiles for the crack initiation phase and 179461 profiles for the crack to propagate to the critical failure length. Hence the lifetime can be estimated by Eqn. 7.9. The temperature cycles within the profile are all based on normal operation, hence they are smaller than that in the accelerated test. Therefore the damage caused by one mission profile is smaller than the threshold and the crack initiation phase exists.

$$Lifetime = 2879 \cdot (74111 + 179461)\ seconds = 730033788\ seconds = 23.1\ years \tag{7.9}$$

Fig. 7.10 shows the characteristics of the nonlinear damage accumulation for the temperature profile in crack propagation phase. The temperature profile consists of 29 cycles with different ΔT_j. Hence the Δr for each cycle are different, as shown in Fig. 7.10(c). The

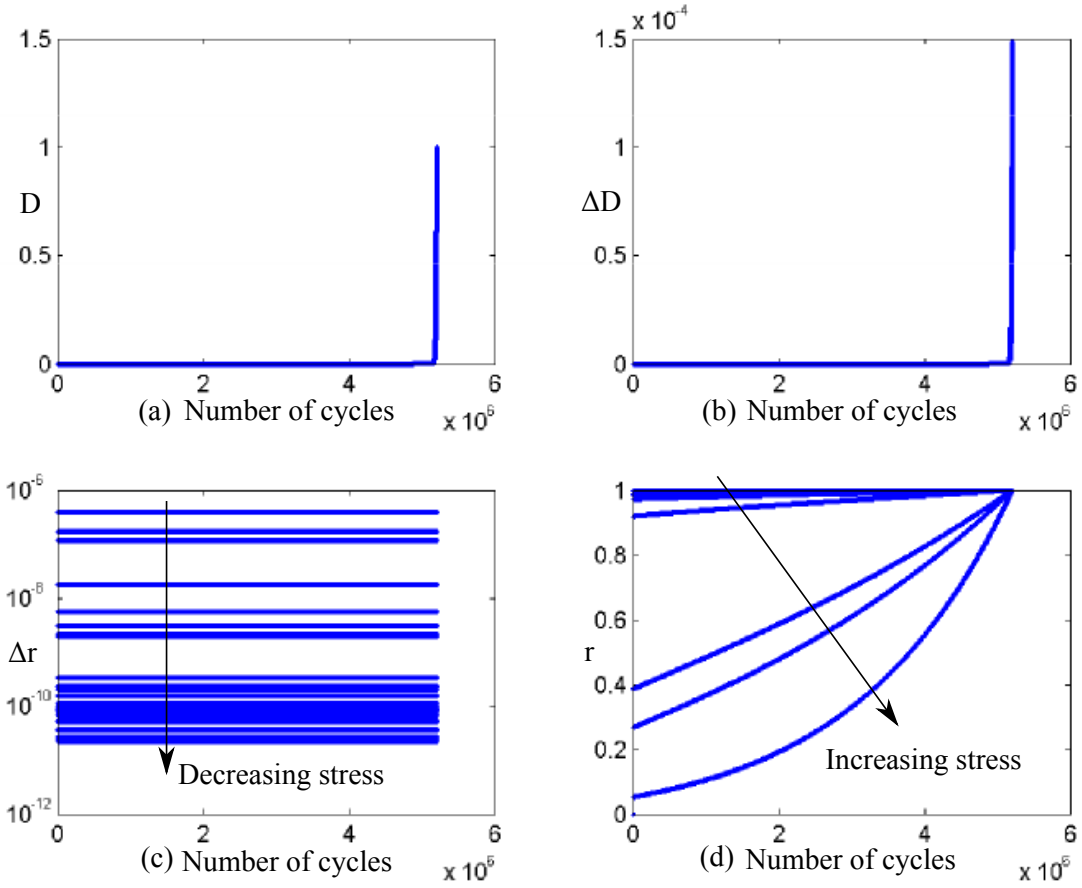


Figure 7.10: Damage accumulation of the temperature profile in crack propagation phase, (a) damage accumulation, (b) damage per cycle, (c) Δr per cycle, (d) variation of r during the lifetime.

D-r curves of 29 cycles are different hence r varies for each cycle and gradually approaches one along different curves, as shown in Fig. 7.10(d). Fig. 7.10(a) and (b) shows the similar property as the two-level stress simulation. The damage accumulates faster as r approaches one, since small temperature cycles could cause serious damage as the large temperature cycles. Fig. 7.11 shows the damage throughout lifetime by different accumulation methods. Clearly, the nonlinear accumulation method gives a more conservative prediction than the linear method for the multi-level temperature profile. This is similar to the result of two-level stress simulation.

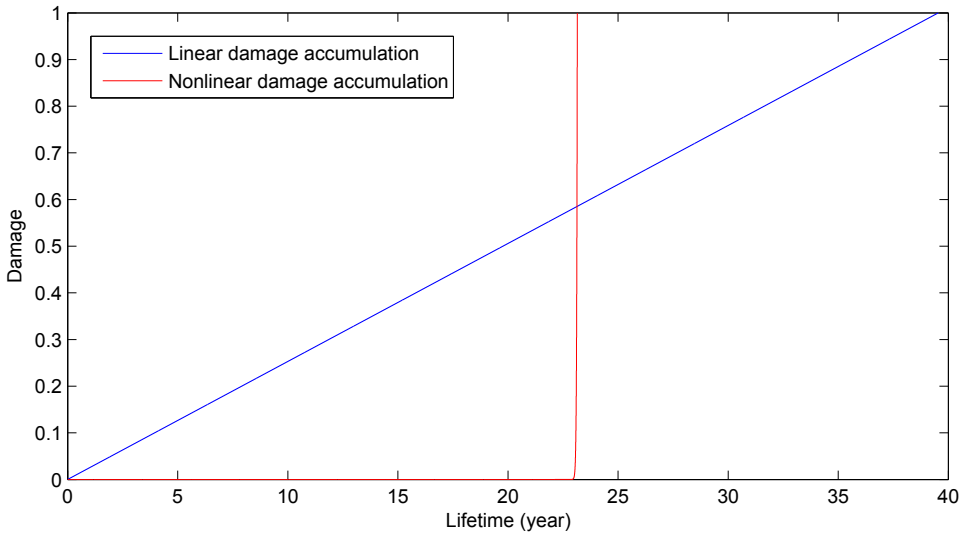


Figure 7.11: Damage accumulation throughout lifetime by the linear and nonlinear methods.

7.2.3 Implementation of the Damage Feedback Loop

As fatigue damage accumulates, the thermal resistance of die-attach solder will increase which leads to a higher junction temperature. This forms a positive feedback and will accelerate the damage accumulation process. Hence ideally, the update of temperature profile is necessary in order to obtain an accurate estimation. The effect of the thermal resistance increase on the damage accumulation depends on how the damage accumulates. As shown in Fig. 7.11, the linear damage accumulation method leads to a large proportion of lifetime with high damage. While for the nonlinear method, the lifetime proportion with high damage is much smaller since the slope of the damage increase is huge once it starts to accumulate. Therefore, the increase of the thermal resistance affects the linear accumulation much more than the nonlinear method.

In order to update of temperature profile, it is necessary to run the electrothermal simulation with increased thermal resistance, which is time consuming. Hence the update is not desirable to be too frequent. 10% increase of the thermal resistance is selected as the

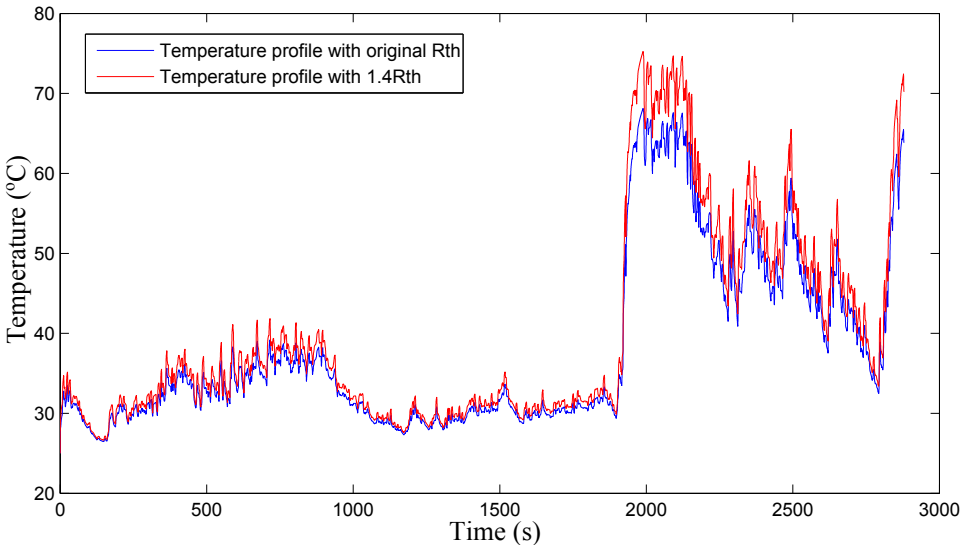


Figure 7.12: Temperature profile with the initial and increased thermal resistance.

threshold. Hence, the inverter electrothermal simulation was run with initial R_{th}, $1.1R_{th}$, $1.2R_{th}$, $1.3R_{th}$, and $1.4R_{th}$ respectively. With the increased thermal resistance, not only the amplitude of temperature cycles were increased but also the number of cycles in the plastic region might rises. Fig. 7.12 shows the temperature profiles with the initial and increased thermal resistance. Table 7.1 shows the detail results of the lifetime estimation with feedback loop. The simulation result agrees with the theoretical analysis, the maximum junction temperature, the amplitude of temperature cycle, and the damage per temperature profile increase with the rise of thermal resistance. Hence the lifetime corresponding to each ΔR_{th} region decreases. Fig. 7.13 shows the comparison of the three damage accumulation methods. The total estimated lifetime of the linear method with feedback loop is 28.1 years which is more conservative than the linear method without feedback loop but less conservative than the nonlinear method. The nonlinear method considers the feedback mechanism (damage accumulates faster when approaching the end of device lifetime) and the lifetime in high damage region is short, hence it is not necessary to apply the feedback loop to the nonlinear method.

Table 7.1: Lifetime estimation with feedback loop. N indicates the number of cycles in one profile and D indicates upper damage limit.

	D	**N**	ΔT_{jmax}	**Damage per profile**	**Number of profile**	**Lifetime (years)**
R_{th}	0.2	29	43.1 ℃	2.3084×10^{-6}	86640	7.9
$1.1R_{th}$	0.4	39	44.9 ℃	2.7912×10^{-6}	71654	6.5
$1.2R_{th}$	0.6	42	46.6 ℃	3.3578×10^{-6}	59563	5.4
$1.3R_{th}$	0.8	47	48.4 ℃	4.0182×10^{-6}	49774	4.5
$1.4R_{th}$	1.0	48	50.2 ℃	4.7830×10^{-6}	41815	3.8

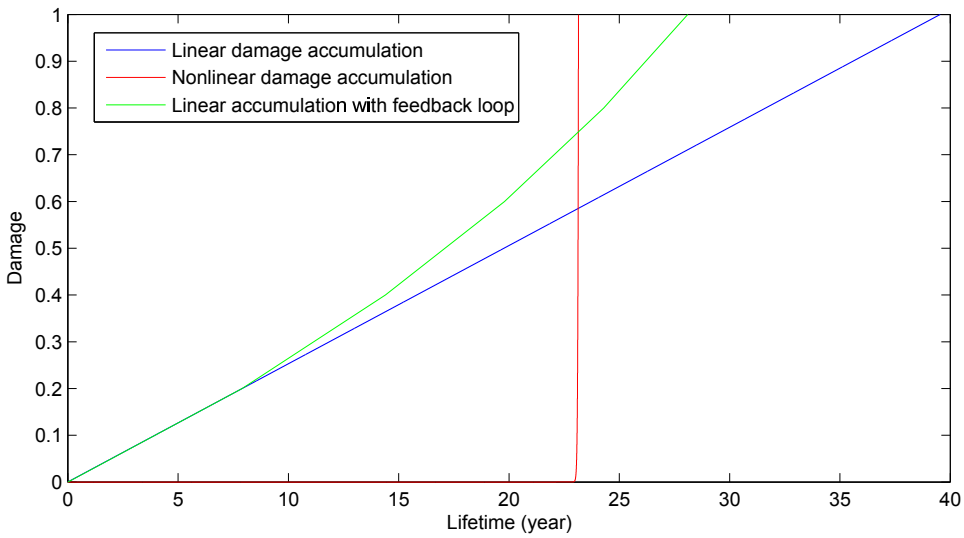


Figure 7.13: Comparison of the estimated lifetime with and without feedback loop.

7.3 Conclusions

The inverter lifetime damage model is applied to a wind power application with a wind speed profile based on real time measurements. The topologies of the wind power generation system and the control strategy of the grid connected inverter were introduced and discussed. The system with fully rated power semiconductors and the predictive current controller was selected for the consideration of simplification and accuracy. The temperature profile generated by the electrothermal simulation was analysed by the rainflow code to obtain the number of valid cycles within one profile and their amplitudes.

Different damage accumulation methods have been used to estimate the inverter lifetime. The nonlinear method gives a more conservative estimation than the linear method since it considers the effect of thermal resistance increase with the rise of accumulated damage. The five-step feedback loop was applied to the linear method and shows a more conservative estimation than without the feedback loop. The nonlinear method shows a better fit with the experimental data and a more conservative lifetime estimation, hence this method is recommended. However, the nonlinear method needs more detailed experimental data to extract x_i which limits its application. The linear method is simpler to apply and widely used, however the implementation of feedback loop is suggested to represent the physical based mechanism (acceleration of damage accumulation) and obtain a more accurate estimation. With this lifetime prediction model, it is possible to investigate the effect of wind profile on inverter lifetime for different applications.

Chapter

8 Conclusions and Future work

8.1 Conclusions

The power generated from different sources of renewable energy is irregular. In the case of wind energy, the variable speed operation gives higher energy conversion rate but it also generates variable frequency AC voltage. For the solar PV system, the generated voltage is DC. This power need to be converted to an AC voltage with fixed frequency in order to connect to the grid. Therefore, a power conversion unit is employed and its reliability is important for the renewable energy systems. The aim of this work is to build an inverter lifetime prediction model with physical meanings to replace the component before the failure occurs. This technique is especially desirable for the systems whose reliability is critical, such as off-shore wind farm. In order to obtain an accurate lifetime estimation, the following techniques are needed: accurate inverter electrothermal model, device reliability data, cycle counting method, and damage accumulation method.

The performance of a power electronic converter depends on the behavior of its semiconductor components. Physics based IGBT and diode models are introduced to obtain the switching losses of the power module at different operating conditions. The model is fit-

ted to a selected IGBT module (SKM50GB123D) according to its datasheet and switching waveforms. The accuracy of this model is shown by comparing the simulation results with the experimental waveforms. The switching power loss LUTs are generated by running the device models at different operating conditions. The IGBT module and heatsink thermal network is extracted from their warming curves and followed by an optimisation procedure to minimise the error between the simulation results and the experimental waveforms. The inverter electrothermal model is built in PLECS which is a tool for high speed simulations of power electronic systems in Matlab/Simulink. The power loss LUTs and the thermal network of IGBT module are stored in a thermal description file and are accessed by PLECS during the simulation for the thermal analysis.

The failure of IGBT modules are mainly caused by the fatigue damage of bond wire and die-attach solder. Fatigue damage is an old but unsolved problem. The current technique of estimating the power device lifetime due to fatigue damage is based on accelerated tests. A power cycling rig is designed and built to obtain the reliability information of the selected IGBT module. Two sets of accelerated test are designed to extract the coefficients of the empirical based fatigue damage model. With the on-state voltage compensation technique, the dominant failure mechanism is believed to be the solder fatigue since the IGBT on-state voltage only starts to increase when its junction temperature increased greatly. From the experimental results, it is clear that the IGBT module lifetime decreases with the increase of ΔT_j. Furthermore, only the crack initiation is greatly temperature dependent, while the propagation of crack is almost independent with ΔT_j. This means the damage caused by large and small ΔT_j is similar when crack starts to propagate.

Different cycle counting methods are discussed and rainflow method is selected since it is able to capture the most characteristics of a mission profile. The temperature profile generated by the inverter electrothermal simulation is then analysed by the rainflow code to count the number of cycles and their amplitudes within this profile.

Commonly used linear damage accumulation method is introduced and its low accuracy is

discussed. A nonlinear accumulation method is presented and shown to fit the experimental data better. The physical meanings of the nonlinear method is discussed. The two-step stress simulation shows that the nonlinear method tends to give a more conservative estimation than the linear damage model. This agrees with the common conclusion that Miner's law always give a less conservative estimation. Also, the simulation result shows that the damage caused by the same ΔT_j gradually increases while the IGBT module is approaching it's lifetime limit. Furthermore, the damage caused by larger ΔT_j is much greater than that caused by smaller ΔT_j at the beginning of its lifetime and this difference gradually decreases as the module approaching it's end of life. This agrees with the experimental results observed from the power cycling test. However, the nonlinear method requires more detailed reliability data which is not always available. A feedback loop is designed to represent the increase of the thermal resistance due to the damage accumulation hence a better estimation could be obtained.

Finally, an application of the lifetime estimation model is implemented in a wind energy generation system. A measurement based wind speed profile is used as the input of the simulation. The electrothermal simulation result shows a reasonable junction temperature variation with the wind speed fluctuation. Three damage accumulation methods are used to estimate the inverter lifetime. The nonlinear method shows the most conservative estimation while the linear method gives the least conservative estimation. Considering both the fitting behavior of the experimental data and the simulation result, the nonlinear method is recommended although it adds some complexities during the accumulation process. When the detailed reliability information is not available, the linear method with feedback loop is recommended to obtain a more accurate estimation.

Table 8.1: The designed test conditions to investigate the effects of both temperature cycle and mean temperature.

	T_{jmin}	T_{jmax}	ΔT_j	T_m	T_{cmax}
Test1	50 ℃	130 ℃	80 ℃	90 ℃	110 ℃
Test2	30 ℃	150 ℃	120 ℃	90 ℃	115 ℃
Test3	30 ℃	110 ℃	80 ℃	70 ℃	95 ℃

8.2 Suggestions for Further Work

8.2.1 Obtain More Power Cycling Data

Due to the time limitation, this work only investigated the main factor that affects the fatigue damage, ΔT_j. However, Held [17] reported that the lifetime is also affected by the mean value of the temperature cycle. At least one more set of test data is required to obtain the lifetime dependence on mean temperature. The test conditions are listed in Table 8.1. This can be estimated by keeping ΔT_j the same as test 1 while decreasing the mean temperature from 90 ℃ to 70 ℃, which is named Test 3 in the table.

The power cycling test under low temperature cycles are desired for the validation of the fatigue damage model since the failure mechanisms for high cycle and low cycle might be different as reported by Scheuerman [57]. More power cycling results could always help improve the accuracy of the fatigue damage model hence give a more accurate lifetime prediction. Furthermore, the power cycling results under low temperature cycles could be used to improve the accuracy of the x_i model, which at this stage is only suitable under high temperature cycles.

Furthermore, the test device can be extended to SiC power MOSFETs. The SiC power devices are developing rapidly and shows a better operation behavior than traditional Si devices. The reliability investigation of such devices is important to boost their application and would be the first time that has ever been carried out.

8.2.2 Inverter Model Validation

The accuracy of the inverter electrothermal model depends on the accuracy of the power loss LUTs and the thermal networks. Although they are proved separately, the implemented inverter model is still not covered by the experimental result directly. Hence an inverter test rig with the selected IGBT power modules could be built to validate its electrothermal model. The IGBT junction temperature needs to be monitored by IC in realtime for a certain mission profile. The same mission profile can be apply to the inverter model and the corresponding temperature profile can be simulated. The accuracy of the inverter model could be obtained by comparing the experimental results with the simulation profile.

8.2.3 Optimum Wind Turbine Control

According to the results of the electrothermal simulation as shown in Fig. 7.7, it is clear that large temperature cycles tend to appear when the wind speed is fast. According to the rainflow damage histogram, the accumulated damage is dominated by the large temperature cycles. However, more power can be generated from wind energy when wind speed is high, as shown in Fig. 7.3. Therefore, the operation strategy of the wind turbine can be optimised by considering both the energy conversion efficiency and reliability to obtain a maximum possible power generation. For example, when the reliability is poor and the time to next possible maintenance is long, the wind turbine could work in a reduced power rate and stable speed to achieve a longer lifetime.

Appendix

A Publication

1. Hui Huang, Angus Bryant, Philip Mawby, "Electrothermal Modelling of Three Phase Inverter", *European Conference on Power Electronics and Applications (EPE)*, Birmingham, UK, August 2011, pp. 1-7.

2. Hui Huang, Philip Mawby, Mike Robert Jennings "Nonlinear Damage Accumulation for Inverter Lifetime Prediction", *EPE Joint Wind Energy and T&D Chapters Seminar*, Aalborg, Denmark, June 2012.

3. Hui Huang, Philip Mawby "Lifetime estimation for grid connected voltage source inverter in wind power generation systems ", *Submitted to IEEE Transactions on Power Electronics*

Appendix

B Data Sheet of Selected IGBT Module

back zurück

SEMIKRON

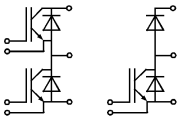
Absolute Maximum Ratings Symbol	Conditions [1]	Values ... 123 D	Units
V_{CES}		1200	V
V_{CGR}	R_{GE} = 20 kΩ	1200	V
I_C	T_{case} = 25/80 °C	50 / 40	A
I_{CM}	T_{case} = 25/80 °C; t_p = 1 ms	100 / 80	A
V_{GES}		± 20	V
P_{tot}	per IGBT, T_{case} = 25 °C	310	W
T_j, (T_{stg})		– 40 . . . +150 (125)	°C
V_{isol}	AC, 1 min.	2 500	V
humidity	DIN 40 040	Class F	
climate	DIN IEC 68 T.1	40/125/56	
Diodes			
I_F = – I_C	T_{case} = 25/80 °C	50 / 40	A
I_{FM} = – I_{CM}	T_{case} = 25/80 °C; t_p = 1 ms	100 / 80	A
I_{FSM}	t_p = 10 ms; sin.; T_j = 150 °C	550	
I^2t	t_p = 10 ms; T_j = 150 °C	1500	A^2s

Characteristics Symbol	Conditions [1]	min.	typ.	max.	Units
$V_{(BR)CES}$	V_{GE} = 0, I_C = 1 mA	≥ V_{CES}	–	–	V
$V_{GE(th)}$	V_{GE} = V_{CE}, I_C = 2 mA	4,5	5,5	6,5	V
I_{CES}	V_{GE} = 0 T_j = 25 °C	–	0,3	1	mA
	V_{CE} = V_{CES} T_j = 125 °C	–	3	–	mA
I_{GES}	V_{GE} = 20 V, V_{CE} = 0	–	–	200	nA
V_{CEsat}	I_C = 40 A V_{GE} = 15 V;	–	2,5(3,1)	3(3,7)	V
V_{CEsat}	I_C = 50 A T_j = 25 (125) °C	–	2,7(3,5)	–	V
g_{fs}	V_{CE} = 20 V, I_C = 40 A		30	–	S
C_{CHC}	per IGBT	–	–	350	pF
C_{ies}	V_{GE} = 0	–	3300	4000	pF
C_{oes}	V_{CE} = 25 V	–	500	600	pF
C_{res}	f = 1 MHz	–	220	300	pF
L_{CE}		–	–	30	nH
$t_{d(on)}$	V_{CC} = 600 V	–	70	–	ns
t_r	V_{GE} = + 15 V / - 15 V [3]	–	60	–	ns
$t_{d(off)}$	I_C = 40 A, ind. load	–	400	–	ns
t_f	R_{Gon} = R_{Goff} = 27 Ω	–	45	–	ns
E_{on} [5]	T_j = 125 °C	–	7	–	mWs
E_{off} [5]		–	4,5	–	mWs
Diodes [8]					
V_F = V_{EC}	I_F = 40 A V_{GE} = 0 V;	–	1,85(1,6)	2,2	V
V_F = V_{EC}	I_F = 50 A T_j = 25 (125) °C	–	2,0(1,8)	–	V
V_{TO}	T_j = 125 °C	–	–	1,2	V
r_T	T_j = 125 °C	–	–	22	mΩ
I_{RRM}	I_F = 40 A; Tj = 25 (125) °C [2]	–	23(35)	–	A
Q_{rr}	I_F = 40 A; T_j = 25 (125) °C [2]	–	2,3(7)	–	µC
Thermal Characteristics					
R_{thjc}	per IGBT	–	–	0,4	°C/W
R_{thjc}	per diode	–	–	0,7	°C/W
R_{thch}	per module	–	–	0,05	°C/W

SEMITRANS® M
IGBT Modules

SKM 50 GB 123 D
SKM 50 GAL 123 D

SEMITRANS 2



GB GAL

Features
- MOS input (voltage controlled)
- N channel, Homogeneous Si
- Low inductance case
- Very low tail current with low temperature dependence
- High short circuit capability, self limiting to 6 * I_{cnom}
- Latch-up free
- Fast & soft inverse CAL diodes [8]
- Isolated copper baseplate using DCB Direct Copper Bonding Technology
- Large clearance (10 mm) and creepage distances (20 mm).

Typical Applications: → B 6 - 85
- Three phase inverter drives
- Switching (not for linear use)

1) T_{case} = 25 °C, unless otherwise specified
2) I_F = – I_C, V_R = 600 V, – di_F/dt = 800 A/µs, V_{GE} = 0 V
3) Use V_{GEoff} = -5 ... -15 V
5) See fig. 2 + 3; R_{Goff} = 27 Ω
8) CAL = Controlled Axial Lifetime Technology.

Case and mech. data → B 6 - 86
SEMITRANS 2

Figure B.1: IGBT module SKM50GB123D datasheet page 1.

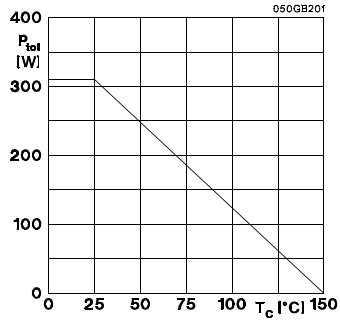
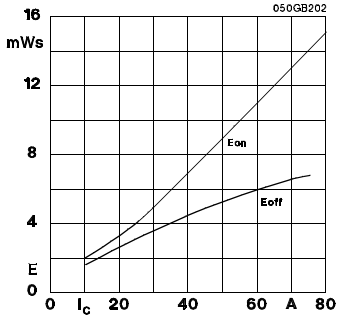
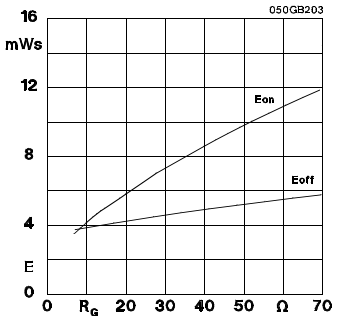


Fig. 1 Rated power dissipation $P_{tot} = f(T_C)$



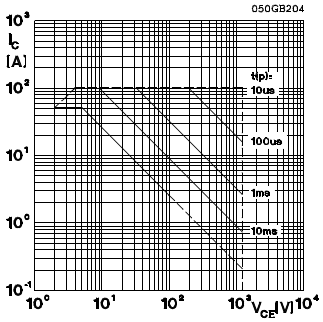
T_j = 125 °C
V_{CE} = 600 V
V_{GE} = ± 15 V
R_G = 27 Ω

Fig. 2 Turn-on /-off energy = f (I_C)



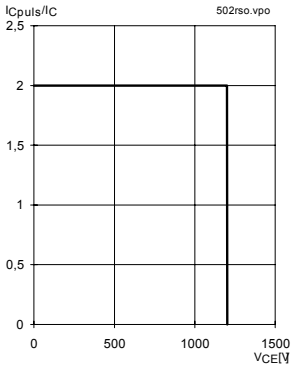
T_j = 125 °C
V_{CE} = 600 V
V_{GE} = ± 15 V
I_C = 40 A

Fig. 3 Turn-on /-off energy = f (R_G)



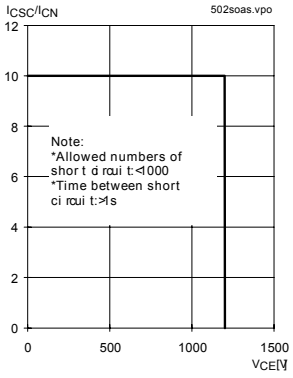
1 pulse
T_C = 25 °C
$T_j \leq$ 150 °C

Fig. 4 Maximum safe operating area (SOA) $I_C = f(V_{CE})$



$T_j \leq$ 150 °C
V_{GE} = ± 15 V
R_{Goff} = 27 Ω
I_C = 40 A

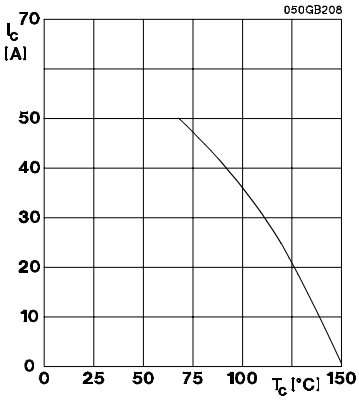
Fig. 5 Turn-off safe operating area (RBSOA)



$T_j \leq$ 150 °C
V_{GE} = ± 15 V
$t_{sc} \leq$ 10 µs
L < 25 nH
I_{CN} = 40 A

Fig. 6 Safe operating area at short circuit $I_C = f(V_{CE})$

Figure B.2: IGBT module SKM50GB123D datasheet page 2.



T_j = 150 °C
$V_{GE} \geq$ 15 V

Fig. 8 Rated current vs. temperature $I_C = f(T_C)$

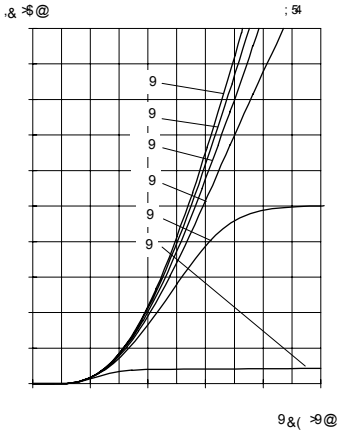


Fig. 9 Typ. output characteristic, t_p = 80 µs; 25 °C

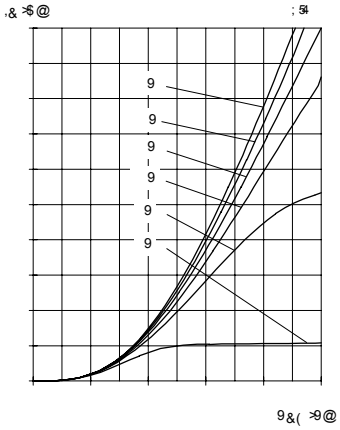


Fig. 10 Typ. output characteristic, t_p = 80 µs; 125 °C

$P_{cond(t)} = V_{CEsat(t)} \cdot I_{C(t)}$

$V_{CEsat(t)} = V_{CE(TO)(Tj)} + r_{CE(Tj)} \cdot I_{C(t)}$

$V_{CE(TO)(Tj)} \leq 1{,}5 + 0{,}002\ (T_j - 25)$ [V]

typ.: $r_{CE(Tj)} = 0{,}02 + 0{,}00008\ (T_j - 25)$ [Ω]

max.: $r_{CE(Tj)} = 0{,}03 + 0{,}00010\ (T_j - 25)$ [Ω]

valid for $V_{GE} = +15^{+2}_{-1}$ V; $I_C > 0{,}3\ I_{Cnom}$

Fig. 11 Saturation characteristic (IGBT)
Calculation elements and equations

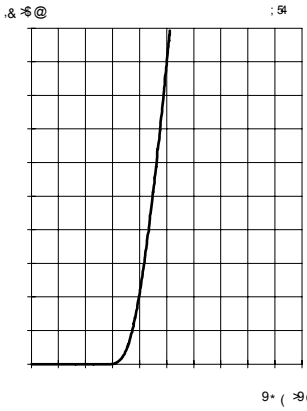


Fig. 12 Typ. transfer characteristic, t_p = 80 µs; V_{CE} = 20 V

Figure B.3: IGBT module SKM50GB123D datasheet page 3.

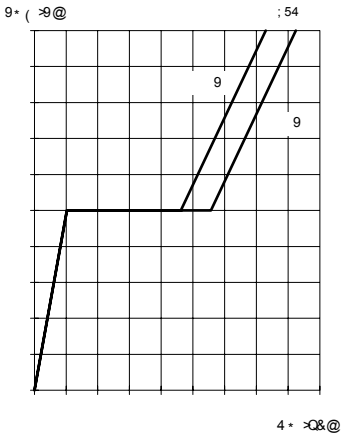
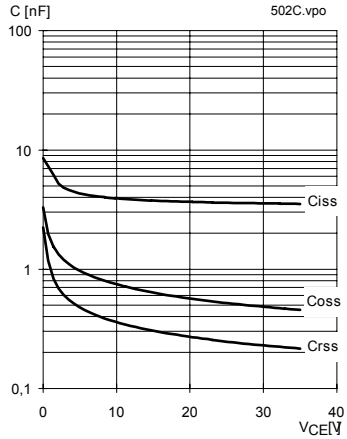


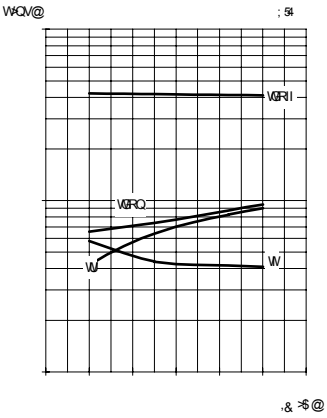
Fig. 13 Typ. gate charge characteristic

I_{Cpuls} = 50 A



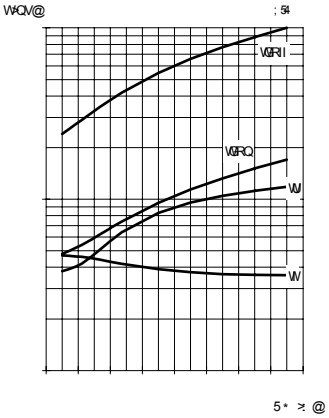
V_{GE} = 0 V
f = 1 MHZ

Fig. 14 Typ. capacitances vs. V_{CE}



T_j = 125 °C
V_{CE} = 600 V
V_{GE} = ± 15 V
R_{Gon} = 27 Ω
R_{Goff} = 27 Ω
induct. load

Fig. 15 Typ. switching times vs. I_C



T_j = 125 °C
V_{CE} = 600 V
V_{GE} = ± 15 V
I_C = 40 A
induct. load

Fig. 16 Typ. switching times vs. gate resistor R_G

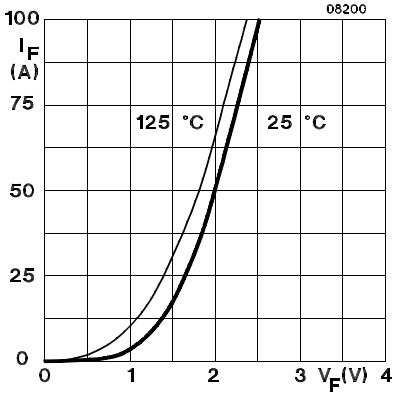


Fig. 17 Typ. CAL diode forward characteristic

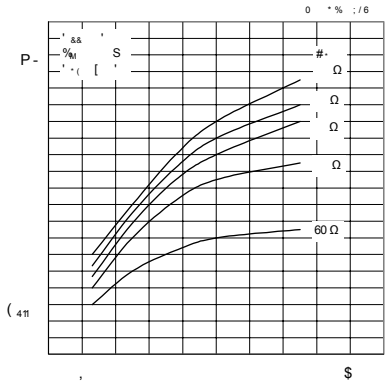


Fig. 18 Diode turn-off energy dissipation per pulse

Figure B.4: IGBT module SKM50GB123D datasheet page 4.

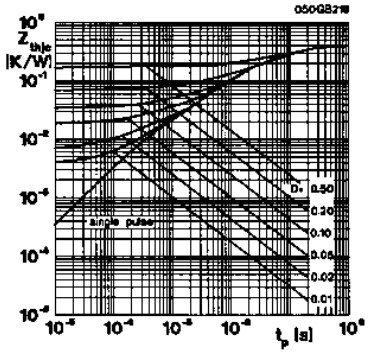


Fig. 19 Transient thermal impedance of IGBT
$Z_{thJC} = f(t_p)$; $D = t_p / t_c = t_p \cdot f$

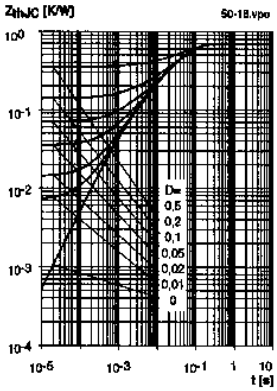


Fig. 20 Transient thermal impedance of inverse CAL diodes $Z_{thJC} = f(t_p)$; $D = t_p / t_c = t_p \cdot f$

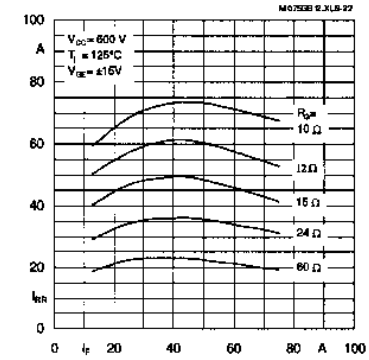


Fig. 22 Typ. CAL diode peak reverse recovery current
$I_{RR} = f(I_F, R_G)$

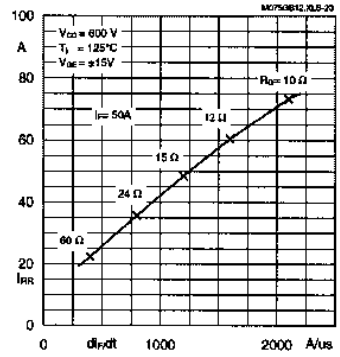


Fig. 23 Typ CAL diode peak reverse recovery current
$I_{RR} = f(di/dt)$

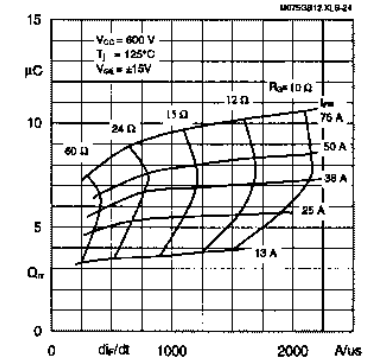
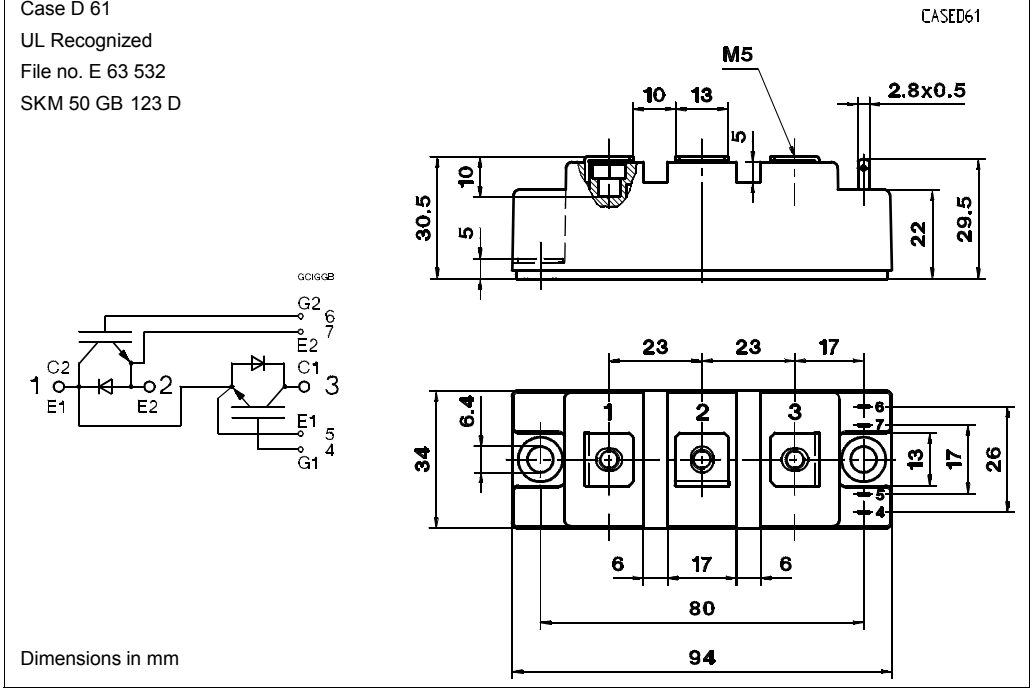


Fig. 24 Typ. CAL diode recovery charge

Typical Applications
include
Switched mode power supplies
DC servo and robot drives
Inverters
DC choppers
AC motor speed control
Inductive heating
UPS Uninterruptable power supplies
General power switching applications
Electronic (also portable) welders
Pulse frequencies also above 15 kHz

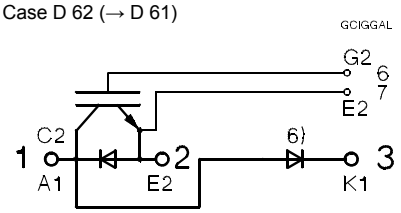
Figure B.5: IGBT module SKM50GB123D datasheet page 5.

SEMITRANS 2
Case D 61
UL Recognized
File no. E 63 532
SKM 50 GB 123 D



Dimensions in mm

SKM 50 GAL 123 D
Case D 62 (→ D 61)



Case outline and circuit diagrams

Mechanical Data						
Symbol	Conditions		Values			Units
			min.	typ.	max.	
M_1	to heatsink, SI Units	(M6)	3	–	5	Nm
	to heatsink, US Units		27	–	44	lb.in.
M_2	for terminals, SI Units	(M5)	2,5	–	5	Nm
	for terminals US Units		22	–	44	lb.in.
a			–	–	5x9,81	m/s^2
w			–	–	160	g

This is an electrostatic discharge sensitive device (ESDS). Please observe the international standard IEC 747-1, Chapter IX.

Eight devices are supplied in one SEMIBOX A without mounting hardware, which can be ordered separately under Ident No. 33321100 (for 10 SEMITRANS 2)
Larger packaging units of 20 or 42 pieces are used if suitable
Accessories → B 6 – 4.
SEMIBOX → C – 1.

Figure B.6: IGBT module SKM50GB123D datasheet page 6.

Appendix

C Matlab Codes

C.1 Find Extreme Points

```
function [ext, exttime] = sig2extforpowercycling(sig, dt)
    % SIG2EXT - searches local extrema from time course (signal),
    %
    % function [ext, exttime] = sig2ext(sig, dt)
    %
    % SYNTAX
    % sig2ext(sig)
    % [ext]=sig2ext(sig)
    % [ext,exttime]=sig2ext(sig)
    % [ext,exttime]=sig2ext(sig, dt)
    %
    % OUTPUT
    % EXT - found extrema from signal SIG,
    % EXTTIME - option, time of extremum occurrence counted from
```

```
% sampling time of DT course (in seconds).
% If no sampling time, DT = 1 is assumed.
%
% INPUT
% SIG - required, time course of loading,
% DT - option, descripion as above, scalar or vector of
% the same length as SIG,
%
% The function caused without an output draws a course graph with
% the searched extrema.
%
error(nargchk(1,2,nargin))
% Output exttime when there are 2 output or no output
TimeAnalize=(nargout==0)|(nargout==2);
% If no sampling time, dt=1 is assumed; else set dt=dt'
if nargin==1,
    dt=1;
else
    dt=dt(:);
end
% sig=sig'
sig=sig(:);
%remove the point which is bigger than previous point and smaller than latter point
w1=diff(sig);
w=logical([1;(w1(1:end-1).*w1(2:end))<=0;1]);
ext=sig(w);
if TimeAnalize,
```

```
    if length(dt)==1,
        exttime=(find(w==1)-1).*dt;
    else
        exttime=dt(w);
    end
end
%remove the ext point which is equal to both previous and next ext point
w1=diff(ext);
w= logical([0; w1(1:end-1)==0 & w1(2:end)==0; 0]);
ext=ext(w);
if TimeAnalize,
    exttime=exttime(w);
end
%remove the ext point which is equal to previous ext point
%for example sig=1 2 3 3 2 1, ext=1 3 1, exttime=0 2.5 5
w= logical([0; ext(1:end-1)==ext(2:end)]);
ext=ext(w);
if TimeAnalize,
    w1=(exttime(2:end)-exttime(1:end-1))./2;
    exttime=[exttime(1:end-1)+w1.* w(2:end); exttime(end)];
    exttime=exttime(w);
end
if length(ext)>2,
    w1=diff(ext);
    w=logical([1; w1(1:end-1).*w1(2:end)<0; 1]);
    ext=ext(w);
    if TimeAnalize,
```

```
        exttime=exttime(w);
    end
end
%remove the point which differences between it and both of its adjacent points are less
%than 1 (or whatever value which is elastic deformation, need calculation)
w1=diff(ext);
w= logical([0;abs(w1(1:end-1)<1).*abs(w1(2:end)<1);0]);
ext=ext(w);
if TimeAnalize,
    exttime=exttime(w);
end
if nargout==0,
    if length(dt)==1,
        dt=(0:length(sig)-1).*dt;
    end
    plot(dt,sig,'b',exttime,ext,'ro')
    legend('SIGNAL','EXTREMA')
    xlabel('time')
    ylabel('signal & extrema')
    clear ext exttime
end
```

C.2 Locate Switching Start/End Point

```
prm_SKM50GB;
    sim_name = 'system_induct_bk';
    sVL=400;
```

```
sIL=50;
sTij=273+130;
sTdj=273+130;
sRg=27;
prm(op_,VL_,base_:current_) = sVL;
prm(op_,IL_,base_:current_) = sIL;
d_T = sTdj; i_T = sTij;
prm(gate_,Rg_,base_:current_) = sRg;
[timeout,xout,yout] = sim(sim_name,[]);
sVak = -yout(:,5);
sVce = yout(:,3);
sIc = yout(:,4);
sVge = yout(:,2);
sIg = yout(:,9);
sVgo = yout(:,10);
sIa = yout(:,6);
sPi = yout(:,11);
sPd = yout(:,12);
sEi = yout(:,7);
sEd = yout(:,8);
% ti1 is the IGBT switch on process starting point
for ti1=1:length(sPi)
    if yout(ti1,1)>=time1
        break
    end
end
% ti3 is the IGBT switch off process starting point
```

```
for ti3=ti1:length(sPi)
    if yout(ti3,1)>=time3
        break
    end
end
for mid1=ti1:ti3
    if yout(mid1,1)>=2.5e-5
        break
    end
end
for mid2=ti3:length(sPi)
    if yout(mid2,1)>=3.9e-5
        break
    end
end
if sPi(mid1)>sPi(mid2)
    ref=mid1;
else
    ref=mid2;
end
% ti2 is the IGBT switch on process finish point
[ext,exttime]=sig2ext(sPi(ti1:mid1),ti1:mid1);
[max_ext,index]=max(ext);
for x=index:2:(length(ext)-1)
    if ext(x)<=1.5*sPi(ref)
        break
    end
```

```
end
ti2=exttime(x);
if ext(x)>1.5*sPi(ref)
    for ti2=exttime(x):ti3
        if sPi(ti2)<=1.5*sPi(ref)
            break
        end
    end
end
figure(1)
plot(yout(exttime,1),ext,'rx')
hold on
% ti4 is the IGBT switch off process finish point
[ext,exttime]=sig2ext(sPi(ti3:mid2),ti3:mid2);
[max_ext,index]=max(ext);
for x=index:2:(length(ext)-1)
    if ext(x)<=1.5*sPi(ref)
        break
    end
end
ti4=exttime(x);
if ext(x)>1.5*sPi(ref)
    for ti4=exttime(x):length(sPi)
        if sPi(ti4)<=1.5*sPi(ref)
            break
        end
    end
```

```
end
plot(yout(exttime,1),ext,'rx')
% td1 is the Diode switch off process starting point
td1=ti1;
% td3 is the Diode switch on process starting point
td3=ti3;
if sPd(mid1)>sPd(mid2)
    ref=mid1;
else
    ref=mid2;
end
% td2 is the Diode switch off process finish point
[ext,exttime]=sig2ext(sPd(td1:mid1),td1:mid1);
[max_ext,index]=max(ext);
for x=index:2:(length(ext)-1)
    if ext(x)<=1.5*sPd(ref)
        break
    end
end
td2=exttime(x);
if ext(x)>1.5*sPd(ref)
    for td2=exttime(x):td3
        if sPd(td2)<=1.5*sPd(ref)
            break
        end
    end
end
```

```
figure(2)
plot(yout(exttime,1),abs(ext),'rx')
hold on
% td4 is the Diode switch on process finish point
[ext,exttime]=sig2ext(sPd(td3:mid2),td3:mid2);
[max_ext,index]=max(ext);
for x=index:2:(length(ext)-1)
    if ext(x)<=1.5*sPd(ref)
        break
    end
end
td4=exttime(x);
if ext(x)>1.5*sPd(ref)
    for td4=exttime(x):length(sPd)
        if sPd(td4)<=1.5*sPd(ref)
            break
        end
    end
end
plot(yout(exttime,1),abs(ext),'rx')
```

C.3 Rainflow Cycle Counting Code

```
figure
    subplot(2,1,1)
    plot(Temperatures(:,1),Temperatures(:,2))
    sig=Temperatures(:,2);
```

```
%Rainflow method
[ext,exttime]=sig2extforpowercycling(sig,Temperatures(:,1)); %find extreme points of signal
subplot(2,1,2)
plot(exttime,ext);
N=length(ext); % Number of extreme values
j=0; i=1;
for index=1:N
    j=j+1;
    %Read next ext point, if there are less than 3 points, read next ext point.
    a(j)=ext(index);
    %if X<Y, read next ext point, if X>=Y, go to next step
    while (j>=3)&&(abs(a(j-1)-a(j-2))<=abs(a(j)-a(j-1)))
        DeltaT=abs(a(j-1)-a(j-2));
        switch(j)
            %if j=3 the while loop is finishing, go to for loop and read new ext point
            case 3
                mean=(a(1)+a(2))/2;
                a(1)=a(2); %discard a(1)
                a(2)=a(3);
                j=2; %read a(3) in the next round
                if DeltaT>4 %discard the cycle which deltaT<4
                    output(1,i)=DeltaT;
                    output(2,i)=mean;
                    output(3,i)=0.5; %count as half cycle
                    i=i+1;
                end
```

```
            otherwise %if j>3 the while loop will continue
                mean=(a(j-1)+a(j-2))/2;
                a(j-2)=a(j);
                j=j-2;
                if DeltaT>4
                    output(1,i)=DeltaT;
                    output(2,i)=mean;
                    output(3,i)=1.0;
                    i=i+1;
                end
        end
    end
end
%Count each range that has not been previously counted as one-half cycle.
for index=1:j-1
    DeltaT=abs(a(index)-a(index+1));
    mean=(a(index)+a(index+1))/2;
    if DeltaT>4
    output(1,i)=DeltaT;
    output(2,i)=mean;
    output(3,i)=0.5;
    i=i+1;
    end
end
halfcycle=find(output(3,:)==0.5);
x=10;
[N1 x]=hist(output(1,:),x);
```

```
if  isempty(halfcycle),
    [N2 x]=hist(output(1,halfcycle),x); % only for half-cycle
    N1=N1-0.5*N2;
end
Damage(1:length(x))=0;
step=x(2)-x(1);
for i=1:length(output(1,:))
    cx=round((output(1,i)-x(1))/step)+1; %calculate which zone the point is in
    if cx>length(x) %the point is in the largest zone
        cx=length(x);
    elseif cx<1 %the point is in the lowest zone
        cx=1;
    end
    Damage(cx)=Damage(cx)+output(3,i)/(0.5*(output(1,i)/871.9)^-4.512);
end
%plot bar chart
figure
subplot(2,1,1)
bar(x,N1,1);
title('Rain Flow Counting Histogram')
xlabel('DeltaT');
ylabel(['N: ' num2str(sum(N1)) ...
        ' (' num2str(sum(N2)/2) ' from half-cycles)']);
subplot(2,1,2)
bar(x,Damage,1);
title('Rain Flow Damage Histogram')
xlabel('DeltaT');
```

```
ylabel('Damage');
sum(Damage)
```

C.4 Nonlinear Damage Accumulation Code

```
a=2.377e13;
b=-4.4457;
c=1.7636e9;
d=-4.2067;
k=1;
UnitR=0;
D1=0;
% Calculate the D and r accumulated in one mission profile
for i=1:length(output(1,:))
    Deltar(i)=output(3,i)/a/output(1,i)^b; % r per cycle
    x(i)=c*output(1,i)^d;
    if i==1
        r(i)=Deltar(i);
    else
        r(i)=D1(i-1)^(1/x(i))+Deltar(i); % accumulated r in one mission profile
    end
    DeltaD(i)=x(i)*(r(i))^(x(i)-1)*Deltar(i);
    D1(i)=sum(DeltaD); % Damage at r
end
% If the accumulated D is smaller than the threshold, it is in crack initiation
% phase, calculate the number of mission profile needed to accumulate
% damage to the threshold.
```

```
threshold=1e-300;
if sum(D1)<threshold
    xmin=c*max(output(1,:))^d;
    r_Initial=threshold^(1/xmin);
    UnitR=sum(Deltar);
    MP_Initial=ceil(r_Initial/UnitR);
    Damage=threshold;
    D2=threshold;
else
    MP_Initial=0;
    Damage=0;
    D2=0;
end
j=2;
MP_Propagation=0;
% Accumulate damage in crack propagation phase.
while(Damage<1)
    for i=1:length(output(1,:))
        Deltar(j)=output(3,i)/a/output(1,i)^b; % r per cycle
        x=c*output(1,i)^d;
        if MP_Initial==0
            R(j)=Deltar(j);
            DeltaD(j)=R(j)^x;
            D2(j)=DeltaD(j);
        else
            R(j)=D2(j-1)^(1/x);
            DeltaD(j)=x*R(j)^(x-1)*Deltar(j);
```

```
            D2(j)=D2(j-1)+DeltaD(j);
        end
        Damage=D2(j);
        j=j+1;
        if Damage>=1
            break
        end
    end
    MP_Propagation=MP_Propagation+1;
end
MP=MP_Initial+MP_Propagation;
Year=2879*MP/3600/24/365;
```

Appendix

D Power Cycling Rig

D.1 Control Flow Chart Diagram of Thermal Network Extraction Test

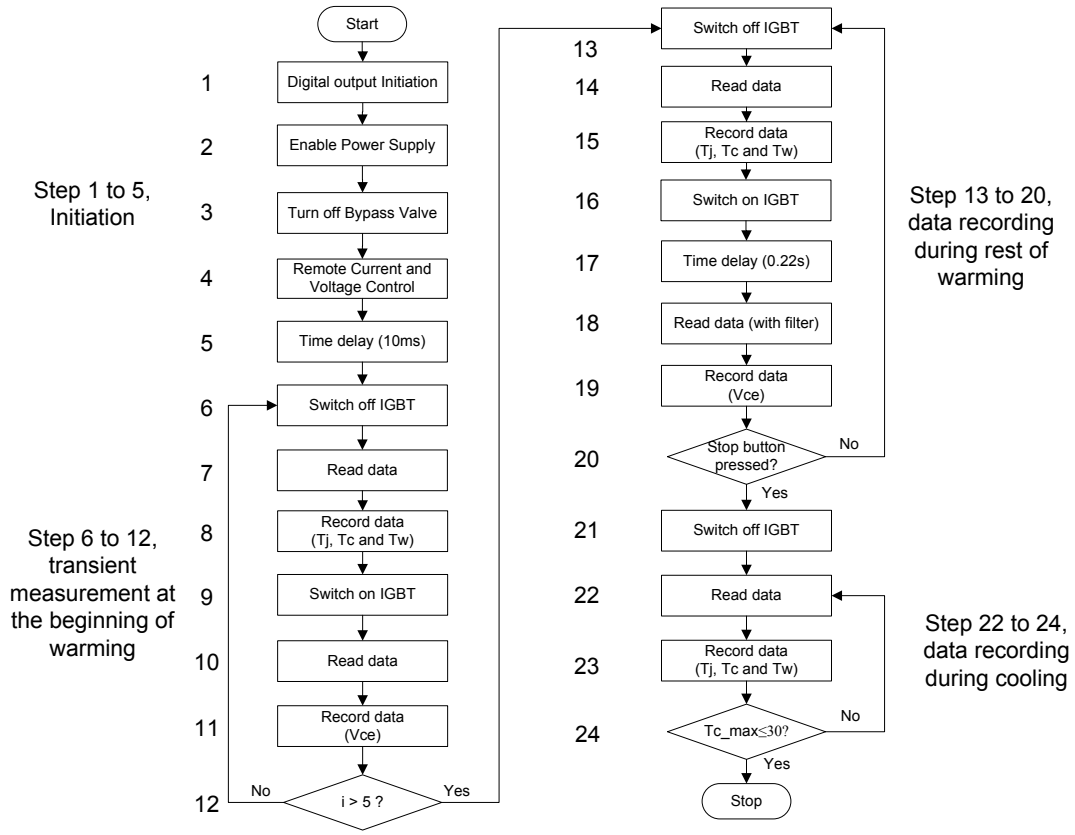


Figure D.1: Control flow chart diagram of thermal network extraction test.

D.2 I/O List of Data Acquisition System

Table D.1: I/O list of data acquisition system

Signals	Signal type	Data type	I/O allocation
V0 - V7	voltage	AI	AI0 - AI7
V8 - V15	voltage	AI	AI32 - AI39
V16 - V22	voltage	AI	AI56 - AI62
Vt1 - Vt8	voltage	AI	AI16 - AI23
Vt9 - Vt14	voltage	AI	AI50 - AI55
Vt15 - Vt22	voltage	AI	AI63 - AI70
Tc1 - Tc8	temperature	AI	AI8 - AI15
Tc9 - Tc16	temperature	AI	AI24 - AI31
Tc17 - Tc24	temperature	AI	AI40 - AI47
Ic	voltage	AI	AI49
Twin	temperature	AI	AI72
Twout	temperature	AI	AI73
Water flow rate	voltage (frequency)	AI	AI48
IGBT gate drive control	boolean	DO	P0.0
Cooling valve control	boolean	DO	P0.1
Bypass valve control	boolean	DO	P0.2
Power supply control	boolean	DO	P0.4
Remote current control	analoge voltage	DO	AO0
Remote voltage control	analoge voltage	DO	AO1

D.3 Circuit Diagram for the DMH

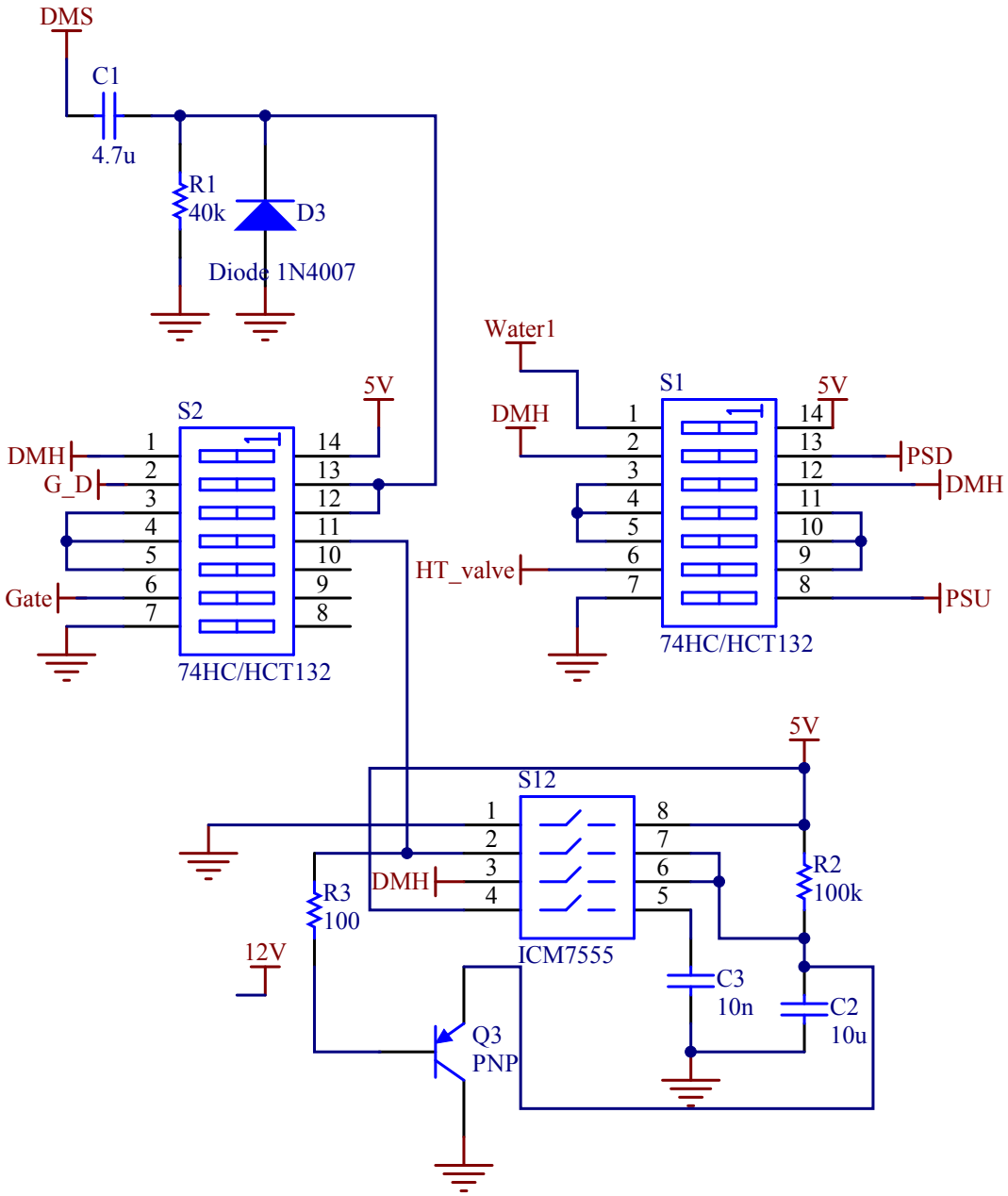


Figure D.2: Circuit diagram of the DMH.

References

[1] *Orthodyne wedge bonders, power ribbon bonding solution.* xiv, 63

[2] W. S. Loh, M. Corfield, H. Lu, S. Hogg, T. Tilford, and C. M. Johnson, "Wire bond reliability for power electronic modules - effect of bonding temperature," in *International Conference on Thermal, Mechanical and Multi-Physics Simulation Experiments in Microelectronics and Micro-Systems*, London, April 2007, pp. 1–6. xiv, 65

[3] M. F. Ashby and D. R. H. Jones, *Engineering Materials 1, an introduction to properties, applications and design.* Butterworth-Heinemann, 2005. xiv, 9, 64, 65, 66, 67, 68, 147

[4] M. Ciappa, "Selected failure mechanisms of modern power modules," *Microelectronics Reliability*, vol. 42, pp. 653–667, 2002. xiv, 10, 67, 68, 72, 125, 129

[5] "Renewables 2010 global status report," REN21, Tech. Rep., 2010. 1, 2

[6] M. Z. Jacobson, "Review of solutions to global warming, air pollution, and energy security," *Energy & Environmental Science*, vol. 2, pp. 148–173, 2008. 1, 2, 3

[7] C. L. Archer and M. Z. Jacobson, "Evaluation of global wind power," *Journal of Geophysical Research*, vol. 110, 2005. 2

[8] R. Singh and B. Baliga, *Cryogenic Operation of Silicon Power Devices.* Kluwer Academic publishers, 1998. 4

[9] P. Lall, M. G. Pecht, and E. B. Hakim, *Influence of temperature on microelectronics and system reliability*. CRC Press, 1997. 6, 73

[10] S. Yang, A. T. Bryant, P. A. Mawby, D. Xiang, L. Ran, and P. Tavner, "An industry-based survey of reliability in power electronic converters," in *Energy Conversion Congress and Exposition (ECCE)*, San Jose, USA, 2009. 6

[11] D. Chamund and D. Newcombe, "Igbt module reliability application notes," Dynex Semiconductor, Tech. Rep., 2010. 8

[12] W. Nelson, *Accelerated Testing Statistical Models, Test Plans and Data Analyses*. A John Wiley & Sons, 2004. 8, 135

[13] H. Lu and C. Bailey, "Lifetime prediction of an igbt power electronics module under cyclic temperature loading conditions," in *Electronic Packaging Technology & High Density Packaging*, Augest 2009. 9

[14] S. Stoyanov, W. Mackay, C. Bailey, D. Jibb, and C. Cregson, "Determination of thermal cross-coupling effects in multi-device power electronic modules," in *International Conference on Power Electronics Machines and Drives*, 2006, pp. 261–265. 9

[15] M. Mermet-Guyennet and M. Piton, "Revisiting power cycling test for better life-time prediction in traction," *Microelectronics Reliability*, vol. 47, pp. 1690–1695, 2007. 10

[16] P. Baginski, "Modules without a baseplate: A reliable and cost-effective solution," *http://www.vincotech.com/news/articles/modules-without-base-plate.html*. 10

[17] M. Held, P. Jacob, G. Nicoletti, P. Scacco, and M. H. Poech, "Fast power cycling test for igbt modules in traction application," in *International Conference on Power Electronics and Drive Systems*, vol. 1, May 1997, pp. 425–430. 10, 62, 110, 120, 126, 129, 139, 183

[18] A. T. Bryant, P. A. Mawby, and P. R. Palmer, "Exploration of power device reliability using compact device models and fast electrothermal simulation," *IEEE Transactions on Industry Applications*, vol. 44, no. 3, pp. 894–903, May/June 2008. 10

[19] D. Hirschmann, D. Tissen, S. Schrder, and R. Doncker, "Reliability prediction for inverters in hybrid electrical vehicles," *IEEE Transactions on Power Electronics*, vol. 22, pp. 2511–2517, November 2007. 10

[20] P. R. Palmer, E. Santi, J. L. Hudgins, X. Kang, J. C. Joyce, and P. Y. Eng, "Circuit simulator models for the diode and igbt with full temperature dependent features," *IEEE Transactions on Power Electronics*, vol. 18, no. 5, pp. 1220–1229, September 2003. 11, 56, 57, 74

[21] X. Kang, A. Caiafa, E. Santi, J. Hudgins, and P. Palmer, "Parameter extraction for a power diode circuit simulator model including temperature dependent effects," in *Applied Power Electronics Conference and Exposition*, Dallas, USA, March 2002. 11

[22] H. Huang, A. Bryant, and P. Mawby, "Electro-thermal modelling of three phase inverter," in *European Conference on Power Electronics and Applications (EPE)*, Birmingham, UK, August 2011, pp. 1–7. 12, 50, 75

[23] S. M. Marco and W. L. Starkey, "A concept of fatigue damage," *Transactions of the ASME*, vol. 76, pp. 627–632, 1954. 13, 147, 148

[24] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, third edition ed. WILEY, 2007. 15, 20, 21, 22, 30, 55

[25] N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics: Converters, Applications and Design*. WILEY, 1989. 16

[26] K. Yamasaki, M. Yoshida, and T. Sugano, "Deep level transient spectroscopy of bulk traps and interface states in si mos diodes," *Japanese Journal of Applied Physics*, vol. 18, no. 1, pp. 113–122, Jan 1979. 17, 19

[27] W. Shockley and J. W. T. Read, "Statistics of the recombinations of holes and electrons," *Phys. Rev.*, vol. 87, pp. 835–842, 1952. 20

[28] S. M. Clemente and D. A. Dapkus, "Igbt models account for switching and conduction losses," in *Power Conversion and Intelligent Motion*, vol. 19, no. 8, Aug. 1993, pp. 51–54. 48

[29] F. Blaabjerg, J. K. Pedersen, S. Sigurjnsson, and A. Elkjaer, "An extended model of power losses in hard-switched igbt-inverters," in *in Proc. Conf. Rec.-IAS Annu. Meeting (IEEE Ind. Applicat. Soc.)*, vol. 3, 1996, pp. 1454–1463. 48

[30] K. E. Atkinson, *An Introduction to Numerical Analysis.* John Wiley & Sons, 1989. 49

[31] J. C. Butcher, *Numerical methods for ordinary differential equations.* John Wiley & Sons, 2003. 49

[32] A. D. Mari, "An accurate numerical steady-state one-dimensional solution of the p-n junction," *Solid-State Electronics*, vol. 11, pp. 1021–1053, 1968. 49

[33] C. L. Ma, P. O. Lauritzen, and P. Y. Lin, "A physically based lumped-charge p-v-n diode model," in *European Conference on Power Electronics and Applications (EPE)*, vol. 2, Brighton, UK, Sep. 1993, pp. 23–28. 49

[34] P. O. Lauritzen, G. K. Andersen, and M. Helsper, "A basic igbt model with easy parameter extraction," in *Power Electronics Specialists Conference (PESC)*, vol. 4, Vancouver, BC, Canada, 2001, pp. 2160–2165. 49

[35] A. R. Hefner, "Analytical modeling of device-circuit interactions for the power insulated gate bipolar transistor," in *in Proc. PESC Rec.IEEE Power Electron. Spec. Conf.*, 1989. 49

[36] A. T. Bryant, P. R. Palmer, E. Santi, and J. L. Hudgins, "Simulation and optimization of diode and insulated gate bipolar transistor interaction in a chopper cell using matlab and simulink," *IEEE Transactions on Industry Applications*, vol. 43, no. 4, pp. 874–883, July/August 2007. 49, 50

[37] A. G. M. Strollo, "Spice modeling of power p-i-n diode using asymptotic waveform evaluation," in *In PESC Conf. Rec.*, Baveno, 1996. 50

[38] B. Allard, H. Morel, and J.-P. Chante, "State-variable modeling of high-level injection regions in power devices," in *Application to power system simulation. In PESC Conf. Rec.*, Toledo, 1992, pp. 885–892. 50

[39] B. Allard, H. Morel, and J. Chante, "Reusing basic semiconductor region models in power device bond graph definition," in *European Conference on Power Electronics and Applications (EPE)*, Brighton, UK, 1993, pp. 34–39. 50

[40] A. T. Bryant, L. Lu, E. Santi, P. R. Palmer, and J. L. Hudgins, "Physical modeling of fast p-i-n diodes with carrier lifetime zoning," *IEEE Transactions on Power Electronics*, vol. 23, no. 1, pp. 189–197, Jan 2008. 50

[41] A. T. Bryant, X. Kang, E. Santi, P. R. Palmer, and J. L. Hudgins, "Two-step parameter extraction procedure with formal optimization for physics-based circuit simulator igbt and p-i-n diode models," *IEEE Transactions on Power Electronics*, vol. 21, no. 2, pp. 295–309, March 2006. 50, 75

[42] M. A. Biondi and S. C. Brown, "Measurements of ambipolar diffusion in helium," *Physical Review*, vol. 75, no. 11, pp. 1700–1705, June 1949. 51

[43] A. T. Bryant, "Simulation and optimisation of diode and igbt interaction in a chopper cell," Ph.D. dissertation, University of Cambridge, 2005. 52, 54

[44] B. J. Baliga, *Power Semiconductor Devices*, Boston, Ed. MA: PWS, 1996. 55

[45] T. Trupke, M. A. Green, P. Wrfel, P. P. Altermatt, and A. Wang, "Temperature dependence of the radiative recombination coefficient of intrinsic crystalline silicon," *Journal of Applied Physics*, vol. 94, pp. 4930–4937, 2003. 57

[46] D. A. Grant and J. Gowar, *Power MOSFETs Theory and Applications.* John Wiley & Sons, 1989. 57

[47] E. R. Motto, "A new low inductance igbt module package," in *PCIM Conference*, Sept. 1996. 60

[48] K. Xing, F. C. Lee, and D. Boroyevich, "Extraction of parasitics within wire-bond igbt modules," in *Applied Power Electronics Conference and Exposition (APEC)*, vol. 1, Feb. 1998, pp. 497–503. 61

[49] M. Trivedi and K. Shenai, "Parasitic extraction methodology for insulated gate bipolar transistors," *IEEE Transactions on Power Electronics*, vol. 15, no. 4, pp. 799–804, Jul. 2000. 61

[50] A. Brambilla, E. Dallago, and R. Romano, "Analysis of an igbt power module," in *Industrial Electronics, Control and Instrumentation (IECON)*, 1994. 61

[51] C. Luechinger, "Large aluminum ribbon bonding - an alternative interconnect solution for power module applications," in *Proc. PCIM*, Nuremberg, 2005. 61, 62

[52] J. Catala, J. Bai, X. Liu, S. Wen, and G.-Q. Lu, "Three dimensional packaging for power semiconductor devices and modules," *IEEE Transactions on Advanced Packaging*, vol. 28, no. 3, pp. 404–412, Aug. 2005. 61

[53] X. Liu, S. Haque, J.Wang, and G.-Q. Lu, "Packaging of integrated power electronics modules using flip-chip technology," *IEEE Transactions on Advanced Packaging*, vol. 1, pp. 290–296, May 2000. 61

[54] P. Beckedahl, T. Grasshoff, and M. Lederer, "A new power module concept for automotive applications," in *Proc. PCIM*, Nuremberg, 2007. 61

[55] C. Gillot, C. Schaeffer, C. Massit, and L. Meysenc, "Double-sided cooling for high power igbt modules using flip chip technology," *IEEE Transactions on Components and Packaging Technologies*, vol. 24, pp. 698–704, Dec. 2001. 61

[56] W. J. Roesch and S. Jittinorasett, "Cycling copper flip chip interconnects," *Microelectronics Reliability*, vol. 44, pp. 1047–1054, 2004. 62

[57] U. Scheuerman and U. Hecht, "Power cycling lifetime of advanced power modules for different temperature swings," in *Proc. PCIM*, Nrnberg, 2002, pp. 59–64. 62, 68, 116, 126, 129, 139, 183

[58] U. Scheuermann and P. Beckedahl, "The road to the next generation power module 100 62

[59] T. Stockmeier, "From packaging to unpackaging - trends in power semiconductor modules," in *Power Semiconductor Devices and IC's (ISPSD)*, May 2008, pp. 12–19. 63

[60] R. Ritchie, "Mechanisms of fatigue-crack propagation in ductile and brittle solids," *International Journal of Fracture*, vol. 100, pp. 55–83, 1999. 64

[61] A. Fatemi and L. Yang, "Cumulative fatigue damage and life prediction theories: A survey of the state of the art for homogeneous materials," *Int. J. Fatigue*, vol. 20, no. 1, pp. 9–34, 1998. 64, 148

[62] S. Ramminger, P. Turkes, and G. Wachutka, "Crack mechanism in wire bonding joints," *Microelectronics Reliability*, vol. 38, pp. 1301–1305, 1998. 66

[63] C. Hager, "Lifetime estimation of al wire bonds based on computational plasticity," Ph.D. dissertation, Swiss Federal Institute of Technology, 2000. 66

[64] W. W. Lee, L. T. Nguyen, and G. S. Selvaduray, "Solder joint fatigue models: review and applicability to chip scale packages," *Microelectronics Reliability*, vol. 40, pp. 231–244, 2000. 69

[65] T. J. Kilinski, J. R. Lesniak, and B. I. Sandor, *Solder joint reliability theory and applications*, J. H. Lau, Ed. New York: Van Nostrand Reinhold, 1991. 69

[66] A. Dasgupta, C. Oyan, D. Barker, and M. Pecht, "Solder creep-fatigue analysis by an energy-partitioning approach," *ASME Journal of Electronic Packaging*, vol. 114, pp. 152–160, 1992. 69

[67] S. X. Wu, J. Chin, T. Grigorich, X. Wu, G. Mui, and C. Yeh, "Reliability analysis for fine pitch bga package," in *Electronic Components and Technology Conference*, June 1998, pp. 737–741. 69

[68] V. Stolkarts, B. Moran, and L. M. Keer, "Constitutive and damage model for solders," in *Electronic Components and Technology Conference*, June 1998, pp. 379–385. 69

[69] K. E. Hokanson and A. Bar-Cohen, "A shear-based optimization of adhesive thickness for die bonding," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 18, pp. 578–584, Sept. 1995. 69

[70] M. Ochiai, T. Akamatsu, and H. Ueda, "Reliability of solder joints assembled with lead-free solder," *Fujitsu scientific and technical journal*, vol. 38, pp. 96–101, 2002. 71

[71] B. Yeung and J. W. Jang, "Correlation between mechanical tensile properties and microstructure of eutectic sn-3.5ag solder," *Journal of Materials Science Letters*, vol. 21, pp. 723–726, 2002. 71

[72] R. Bayerer, T. Hermann, T. Licht, J. Lutz, and M. Feller, "Model for power cycling lifetime of igbt modules various factors influencing lifetime," in *in Proc. of the 5th International Conference on Integrated Power Electronics Systems (CIPS)*, Nuremberg, Germany, March 2008. 71, 125

[73] A. Hamidi, A. Stuck, N. Beck, and R. Zehringer, "Time dependent thermal fatigue of hv-igbt-modules," in *In Proc. of the 27th Kolloquium Halbleiter-Leistungsbauelemente und Materialgte von Silizium*, Oct. 1998. 71

[74] G. Yang, *Life Cycle Reliability Engineering.* John Wiley & sons, 2007. 72

[75] T. B. Lim, "The impact of wafer back surface finish on chip strength," in *IEEE Proceedings of the 27th. Annual Reliability Physics Symposium*, 1989, pp. 131–136. 72

[76] J. S. Suehk, P. Chaparala, C. Messick, and W. M. Miller, "Field and temperature acceleration of time-dependent dielectric breakdown in intrinsic thin sio2," in *IEEE International 32nd Annual Proceedings on Reliability Physics Symposium*, 1994. 73

[77] M. M. Maranowski and J. A. Cooper, "Time-dependent-dielectric-breakdown measurements of thermal oxides on n-type 6h-sic," *IEEE Transactions on Electron Devices*, vol. 46, pp. 520–524, 1999. 73

[78] L. Yu, K. P. Cheung, J. Campbell, J. S. Suehle, and K. Sheng, "Oxide reliability of sic mos devices," in *IEEE International on Integrated Reliability Workshop Final Report*, Oct. 2008, pp. 141–144. 73

[79] X. Kang, A. Caiafa, E. Santi, J. L. Hudgins, and P. R. Palmer, "Characterization and modeling of high-voltage field-stop igbts," *IEEE Transactions on Industry Applications*, vol. 39, no. 4, pp. 922–928, July/August 2003. 74

[80] R. F. Pierret, *Semiconductor Device Fundamentals*, K. Harutunian, Ed. Addison-Wesley Publishing Company, 1996. 79

[81] P. Nance and M. Mrz, "Thermal modelling of power electronic systems," in *European Power Electronics System*, 2000. 90, 91, 92

[82] P. Cova and F. Fantini, "On the effect of power cycling stress on igbt modules," *Microelectronics Reliability*, vol. 38, pp. 1347–1352, 1998. 93

[83] J. Lai and F. Peng, "Multilevel converters - a new breed of power converters," in *Conference Record of the IEEE-IAS Annual Meeting*, 1995, pp. 2348–2356. 99

[84] M. Manjrekar and G. Venkataramanan, "Advanced topologies and modulation strategies for multilevel inverters," in *Conference Record of the IEEE-PESC*, 1996, pp. 1013–1018. 99

[85] M. A. Jabbar, A. M. Khambadkone, and Y. Zhang, "Space-vector modulation in a two-phase induction motor drive for constant-power operation," *IEEE Transactions on Industrial Electronics*, vol. 51, no. 5, pp. 1081–1088, October 2004. 100

[86] B. P. McGrath, D. G. Holmes, and T. Meynard, "Reduced pwm harmonic distortion for multilevel inverters operating over a wide modulation range," *IEEE Transactions on Power Electronics*, vol. 21, no. 4, pp. 941–949, July 2006. 100

[87] L. Yan, X. Li, H. Hu, and B. Zhang, "Research on svpwm inverter technology in wind power generation system," in *Electrical and Control Engineering (ICECE)*, Yichang, China, September 2011, pp. 1220–1223. 103, 105

[88] D. L. Blackburn, "Temperature measurements of semiconductor devices - a review," in *Semiconductor Thermal Measurement and Management Symposium*, March 2004, pp. 70–80. 120

[89] M. Matsuishi and T. Endo, "Fatigue of metals subjected to varying stress," in *Jap Soc. Mechanical Engineering*, 1968. 142

[90] *Standard Practices for Cycle Counting in Fatigue Analysis*, ASTM international Std. 144

[91] W. Hwang and K. S. Han, "Cumulative damage models and multi-stress fatigue life prediction," *Journal of Composite Materials*, vol. 20, pp. 125–153, 1986. 147

[92] F. R. Shanley, "A theory of fatigue based on unbonding during reversed slip," The Rand Corporation, Tech. Rep., 1952. 151

[93] I. Swan, "3-d thermal modelling of power device packaging," Ph.D. dissertation, School of Engineering, University of Warwick, 2010. 158

[94] S. Achilles and M. Pller, "Direct drive synchronous machine models for stability assessment of wind farms," in *Fourth Int. Workshop on Large Scale Integration of Wind Power and Transmission Networks for Offshore Windfarms*, Billund, Denmark, Oct. 2003. 160

[95] F. Blaabjerg, Z. Chen, and S. B. Kjaer, "Power electronics as efficient interface in dispersed power generation systems," *IEEE Transactions on Power Electronics*, vol. 19, no. 5, pp. 1184–1194, 2004. 161

[96] A. B. Plunkett, "A current controlled pwm transistor inverted drive," in *In Proc. IEEE IAS Annu. Meeting*, 1979, pp. 785–792. 165

[97] A. Schonung and H. Stemmler, *Static frequency changers with 'subharmonic' control in conjunction with reversible variable-speed a.c. drives.* Brown Boveri Rev., Aug./Sept. 1964. 165

[98] D. M. Brod and D. W. Novotny, "Current control of vsi-pwm inverters," *IEEE Transactions on Industry Applications*, vol. IA-21, no. 4, pp. 562–570, July/August 1985. 165

[99] D. G. Holmes and D. A. Martin, "Implementation of a direct digital predictive current controller for single and three phase voltage source inverters," in *Annual Meeting IEEE Industry Application*, 1996, pp. 906–913. 166, 169