UNIVERSITY OF CALIFORNIA,
IRVINE

Dynamic Thermal Management of Transistors using Holey Silicon-Based Lateral Thermoelectric Cooler

THESIS

submitted in partial satisfaction of the requirements
for the degree of

MASTER OF SCIENCE

in Mechanical and Aerospace Engineering

by

Jiajian Luo

Thesis Committee:
Associate Professor Jaeho Lee, Chair
Professor Yun Wang
Assistant Professor Camilo Velez Cuervo

2023

DEDICATION

To

my girlfriend and my parents,

in recognition of their love and support

TABLE OF CONTENTS

Page

LIST OF FIGURES v

LIST OF TABLES viii

LIST OF SYMBOLS ix

ACKNOWLEDGEMENTS xi

ABSTRACT OF THE THESIS xii

Chapter 1 Introduction 1
1.1 Microprocessor Development and its Thermal Management Challenges 1
1.2 Self-heating Effect in Transistors 2
1.3 Overview of Conventional Cooling Methods for Microelectronic Devices 4
1.4 Fundamentals of Thermoelectric Cooling 6
1.5 Development of Thermoelectric Coolers in Microelectronic Systems 9
1.7 Vertical TEC vs. Lateral (Planar) TEC 11
1.8 Holey Silicon as a Promising Thermoelectric Material 13
1.9 Research Goals and Objectives 15
Chapter 2 Basic Model Setup 16
2.1 Device Configuration 16
2.1.1 Transistor configurations 16
2.1.2 TEC configurations 17
2.2 Material Selection and Their Properties 18
2.2.1 Thermal conductivity of holey silicon 18
2.2.2 Electrical conductivity and Seebeck coefficient of holey silicon 21
2.2.3 Other material selection and their properties 24
2.3 Other Important Conditions 25
2.3.1 Heating conditions (transient study only) 25
2.3.2 Boundary conditions 27

Chapter 3 Steady-state TEC Cooling 29
3.1 2D Simulation (parametric study) 30
3.1.1 Model selection 30
3.1.2 Reference simulation 31
3.1.3 Effect of hotspot distance 33
3.1.4 Effect of TEC length 35
3.1.5 Effect of ambient temperature 36
3.1.6 Effect of hotspot heat flux 37
3.1.7 Summary of 2D parametric study 39
3.2 3D Simulation 40
3.2.1 Reference simulation 40
3.2.2 Multi-parameter optimization with fixed chip size 43
3.2.3 Multi-parameter optimization with fixed a/b ratio 45
3.3 1D Analytical Model and TEC Advantage 46
Chapter 4 Transient TEC Cooling 51
4.1 Hotspot with Constant Heat Flux 52
4.1.1 Constant current + transient current pulse 54
4.1.2 Single transient current pulse 54
4.2 Hotspot with Transient Heat Flux 56
4.2.1 Reference simulation 56
4.2.2 Amplitude-dependent cooling performance 57
4.2.3 Location-dependent cooling performance 59
4.2.4 Duration-dependent cooling performance 61
4.2.5 Shape-dependent cooling performance 62
Chapter 5 TEC Array Performance 65
5.1 Lateral TEC Array Designs 66
5.2 Array Design Optimization 67
5.2.1 Reference simulation 67
5.2.2 Effect of boundary area 70
5.2.3 Effect of internal distance 72
5.3 Preliminary Study of Dynamic Hotspot 74
Chapter 6 Conclusion and Outlooks 77
References 80

LIST OF FIGURES

Page

Figure 1-1 Microprocessors trend data in 52 years (1971-2023) 2

Figure 1-2 Self-heating effect in SOI power transistors 3

Figure 1-3 Category of conventional cooling methods 5

Figure 1-4 Schematic of a thermoelectric cooler system 6

Figure 1-5 TEC supercooling effect 8

Figure 1-6 Development of thermoelectric coolers 9

Figure 1-7 Holey silicon-based TECs in vertical and lateral configurations 11

Figure 1-8 Holey silicon nanostructures 13

Figure 2-1 Transformation from a specific LDMOS device to a generalized model 16

Figure 2-2 Schematic of a transistor-TEC system using lateral holey silicon-based TEC 17

Figure 2-3 The cross-plane and in-plane thermal conductivity of holey silicon calculated by the spectral scaling model 19

Figure 2-4 Electric transport computation of bulk silicon using ab initio theory. 23

Figure 2-5 Single heat pulse with local background temperature 25

Figure 2-6 Consecutive heat pulse starting with room temperature 25

Figure 2-7 Single heat pulse with local constant heat flux 26

Figure 3-1 Schematic of 2D LDMOS-TEC system using boundary (line) hotspot 30

Figure 3-2 Schematic of 2D LDMOS-TEC system using area hotspot 30

Figure 3-3 Relationship between hotspot temperature and TEC current 31

Figure 3-4 Comparison of temperature distribution between TEC OFF and TEC ON with optimal current 32

Figure 3-5 Relationship between temperature reduction and hotspot distance 33

Figure 3-6 Relationship between temperature reduction and TEC length 35

Figure 3-7 Relationship between temperature reduction and ambient temperature 36

Figure 3-8 Relationship between temperature reduction and hotspot heat flux 37

Figure 3-9 Schematic of the 3D transistor-TEC model 39

Figure 3-10 Relationship between hotspot temperature and TEC current 40

Figure 3-11 Comparison of temperature distribution between TEC OFF and TEC ON with optimal current in 3D reference simulation 41

Figure 3-12 Optimization results with fixed chip size 43

Figure 3-13 Optimization results with fixed a/b ratio 45

Figure 3-14 1D heat transfer model with single passive cooling 46

Figure 3-15 1D heat transfer model with TEC cooling 47

Figure 3-16 TEC Advantage zone 48

Figure 4-1 Schematics of transient TEC model 51

Figure 4-2 Two patterns of transient current and their corresponding cooling results 52

Figure 4-3 Peak hotspot temperature in early supercooling and sequential temperature overshoot 54

Figure 4-4 Concept of using transient TEC cooling to deal with transient heat flux 56

Figure 4-5 Evolution of hotspot temperature under various TEC current amplitudes 57

Figure 4-6 Evolution of hotspot temperature under various temporal locations of TEC current pulse 59

Figure 4-7 Evolution of hotspot temperature under various temporal locations of TEC current 60

Figure 4-8 Hotspot temperature as a function of time under different TEC current pulse shapes 62

Figure 5-1 Thermal image of LDMOS array under different background temperatures 64

Figure 5-2 Three types of lateral TEC array designs 65

Figure 5-3 Geometry of three different designs in reference simulation 67

Figure 5-4 Results of steady-state reference simulation 68

Figure 5-5 Optimization with increased TEC boundary lengths 69

Figure 5-6 The hotspot temperature and the corresponding temperature reduction data in three array designs with increased TEC boundary lengths 70

Figure 5-7 Optimization with increasing internal distance 71

Figure 5-8 The hotspot temperature and the corresponding temperature reduction data in three array designs with increasing internal distance 72

Figure 5-9 Schematics of dynamic hotspot in three array designs 73

Figure 5-10 Temperature distribution maps of three array designs under dynamic hotspot conditions 74

Figure 5-11 Hotspot temperature as a function of the TEC current in four hotspot locations 75

LIST OF TABLES

Page

Table 3-1 Specific parameters used for 2D reference simulation 31

Table 3-2 Specific parameters used for hotspot distance study 33

Table 3-3 Specific parameters used for TEC length study 34

Table 3-4 Specific parameters used for ambient temperature study 35

Table 3-5 Specific parameters used for heat flux study 37

Table 3-6 Summary of parametric study 38

Table 3-7 Specific parameters used for 3D reference simulation 40

Table 3-8 Specific parameters used in the optimization with fixed chip size 42

Table 3-9 Specific parameters used in the optimization with fixed a/b ratio 44

Table 4-1 Specific parameters used in the transient simulation with constant heat flux 52

Table 4-2 Specific parameters used in the transient simulation with time-dependent heat flux 56

Table 4-3 Peak hotspot temperature under different TEC current amplitudes 58

Table 4-4 Peak hotspot temperature under different temporal locations of the TEC current pulse 59

Table 4-5 Peak hotspot temperature under different durations of the TEC current pulse 61

Table 4-6 Summary of the reduction in peak hotspot temperature and the specific TEC current pulse conditions 63

Table 5-1 Specific parameters used in the array design simulation 67

Table 5-2 Specific parameters used in the array simulation with dynamic hotspot conditions 74

LIST OF SYMBOLS

A	activity factor
a_l	lattice spacing
a	hotspot distance
b	TEC length
C	capacitance
c	length (x) of cooler
C_p	specific heat
$C_{V,j}$	phonon specific heat per unit volume
d	length (x) of heater
d_{eff}	characteristic diameter
e	thickness (z) of chip
E	electric field strength vector
E_c	conductance band edge
E_f	Fermi level
f	frequency
f	distance between hotspot and top surface
F(δ)	reduction function
f_0	Fermi-Dirac distribution function
h	Planck constant
ħ	reduced Planck constant
h	convection coefficient
h_{eff}	effective convection coefficient
I	current
J	current density vector
K	thermal conductance
k	thermal conductivity
k_B	Boltzmann constant
$k_{HS,x}$	in-plane thermal conductivity of holey silicon
$k_{HS,z}$	cross-plane thermal conductivity of holey silicon
k_x, k_y	in-plane thermal conductivity
k_z	cross-plane thermal conductivity
L	length (x)
L_{HS}	length (x) of hotspot
m_0	mass of electrons
m_n*	effective mass for electrons
n	neck size
n_0	electron carrier concentration
N(E)	density of states of conductance band for p-type semiconductor/ density of states of valence band for n-type semiconductor
N_C	effective conductance density of states

N_d	donor concentration
N_i	intrinsic carrier concentration
P	power dissipation
p	pitch size
q	heat flux vector
Q, q’	hotspot heat flux
q_0	elementary positive charge
q_c	cooling power
Q_e	internal heat source
R	electrical resistance
R_{th}	thermal resistance
S	Seebeck coefficient
S_n	Seebeck coefficient of n type
S_p	Seebeck coefficient of p type
S_{si}	Seebeck coefficient of (holey) silicon
T	absolute temperature
T_1, T_c, T_{hs}	absolute temperature at the cold end (heat source)
T_2, T_h, T_{surf}	absolute temperature at the hot end (heat sink)
T_∞	ambient temperature
T_{HS}	thickness (z) of hotspot
t	time
V	voltage
v	electron group velocity
v_j	phonon group velocity
x_ω	non-dimensional phonon frequency
θ_j	Debye temperature
$\lambda(E)$	energy-dependent electron MFP
$\lambda_{(z,\infty)}(\omega,n)$	phonon MFP of silicon nanowire
λ_0	electron MFP
$\lambda_1(\omega)$	frequency dependent phonon MFP
λ_{II}	(electron) ionic impurities scattering MFP
λ_{IV}	(electron) intervalley scattering MFP
λ_B	(phonon) boundary scattering MFP
λ_{BS}	(electron) boundary scattering MFP
λ_D	(phonon) point-defect scattering MFP
$\lambda_{e\text{-}p}$	electron-phonon scattering MFP
λ_{opt}	(phonon) optical scattering MFP
λ_{PS}	(electron) plasma scattering MFP
λ_U	(phonon) Umklapp scattering MFP
ρ	density
σ	electrical conductivity
τ_j	phonon relaxation time
ω_D	Debye cut-off frequency

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to Professor Jaeho Lee, my committee chair. He has not only been a mentor in my academic journey but also an exemplary role model in research, collaboration, leadership, and fatherhood.

I am grateful to my sponsors and collaborators, Jingjing Chen and Archana Venugopal from Texas Instruments Incorporated, for their generous financial support and invaluable technical guidance. Without their assistance, I would not have been able to make such significant progress.

I would also like to express my gratitude to Professor Yun Wang and Professor Camilo Velez Cuervo, who have provided me with invaluable support and guidance throughout my M.S. and Ph.D programs. Their exceptional teaching has greatly influenced my development as a skilled engineer and researcher.

I am deeply grateful to my current labmates Jiahui Cao, Hoyeon Park, and Charles Owens for their invaluable help and companionship. Additionally, I would like to express my sincere appreciation to my previous colleagues Jungyun Lim, Zongqing Ren, Ziqi Yu, Jonathan Sullivan, Xiao Nie, and Laia Ferrer-Argemi for their exceptional mentoring and assistance.

I would like to express my sincere gratitude to my family for their unwavering love, support, and encouragement throughout my studies in the United States. They have empowered me to overcome every challenging and stressful moment.

Finally, I would like to express my deepest gratitude to my girlfriend, Shuo Tao, for her unwavering support and boundless love. Her infectious optimism has had a profound impact on me, transforming my perspective towards challenges and stress. She has consistently been there for me whenever I faced difficulties recently.

ABSTRACT OF THE THESIS

Dynamic Thermal Management of Transistors using Holey Silicon-based Lateral Thermoelectric Cooler

by

Jiajian Luo

Master of Science in Mechanical and Aerospace Engineering

University of California, Irvine, 2023

Professor Jaeho Lee, Chair

While the state-of-the-art commercial microelectronic device has achieved a technology node as small as 5nm and even below 5nm, most conventional thermal management techniques are limited to the macroscale level. The serious self-heating effect in micro- and nano-scale transistors necessitates the development of advanced cooling techniques to address the increasingly demanding yet delicate thermal management challenges. In this thesis, we present a novel lateral thermoelectric cooler based on holey silicon that enables dynamic thermal management in microelectronic systems. This holey silicon-based lateral TEC exhibits exceptional TEC cooling performance and is compatible with the conventional microfabrication process, allowing for direct integration into chip architecture and creating an all-in-one package system.

In the subsequent chapters, a comprehensive investigation of the transistor-TEC system is conducted through 2D and 3D simulations using the COMSOL Multiphysics platform. The power transistor (LDMOS) is considered as the cooling object to explore the feasibility of utilizing holey silicon-based lateral TECs for dynamic thermal management in both spatial and temporal domains.

Initially, a parametric study under steady-state conditions examines the influence of various geometrical parameters, boundary conditions, and material properties on TEC cooling performance. Subsequently, a transient TEC analysis investigates its cooling capabilities when employing transient TEC current pulses for managing constant heat fluxes and transient heat pulses. Finally, a comparative study of three different array designs is performed to evaluate their effectiveness in achieving efficient spatial thermal management. In summary, this thesis demonstrates significant advancements in holey silicon-based lateral TECs' cooling performance while showcasing their potential for dynamic thermal management in space and time. Additionally, it provides valuable insights into optimizing the TEC design and lays the foundation for future fabrication and experimental endeavors.

Chapter 1

Introduction

1.1 Microprocessor Development and its Thermal Management Challenges

Over the past five decades, Moore's law has been held as an empirical law in the semiconductor industry. Moore's law posits that the number of transistors on microprocessors doubles every two years, which becomes the industrial guide of feature size reduction and performance improvements in semiconductor devices [1]. Recently in May 2023, Apple released its top-end desktop CPU, M2 Ultra, which incorporates 134 billion transistors in a chip using TSMC 5nm fabrication process [2], again, claiming the successful extension of Moore's law. However, the relentless miniaturization of transistors and the continued growth in their density have brought about significant thermal management challenges in microprocessors, as power dissipation caused by the self-heating effect is inevitable in transistors.

Power dissipation, for many years, has been the primary design constraint for microelectronic systems. In microprocessors, the dominant power dissipation can be attributed to dynamic power, which is given by [3]:

$$P = CV^2Af \tag{1}$$

Here, P is the power dissipation, C is the aggregate load capacitance, V is the supply voltage, A is the activity factor and f is the operating frequency. Although supply voltage has shown significant

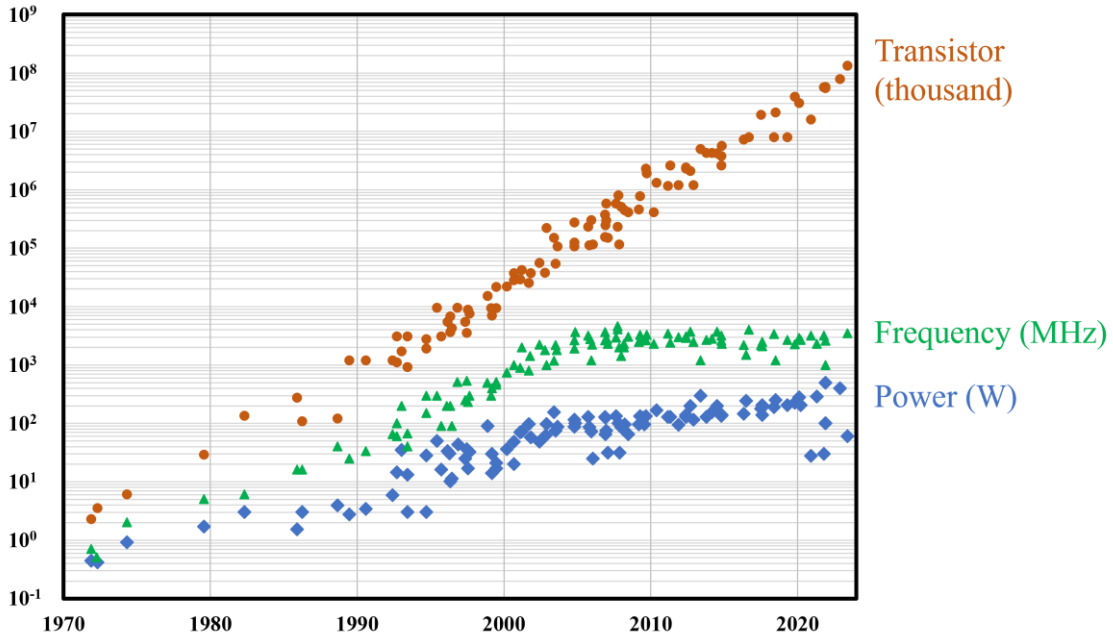

Figure 1-1. Microprocessors trend data in 52 years (1971-2023). Data of 1971-2021adapted from [73].

reduction through transistor scaling, the high operating frequency acts as the major factor for excessive power dissipation. Such high power dissipation within limited chip space generates local hotspots, resulting in excessive chip temperature.

Figure 1-1 shows the trend of microprocessors over 52 years. Until 2023, although the transistor count still follows the exponential growth line, the frequency and power dissipation have reached a bottleneck on the orders of 1 GHz and 100 W.

1.2 Self-heating Effect in Transistors

Transistors are the basic components in microprocessors and, more generally, integrated circuits (ICs). For most digital and some analog applications, metal-oxide-semiconductor field-effect transistors (MOSFETs) have become the major option due to their lower power consumption and high input impedance. Despite minimal current flowing through MOSFETs, intense current

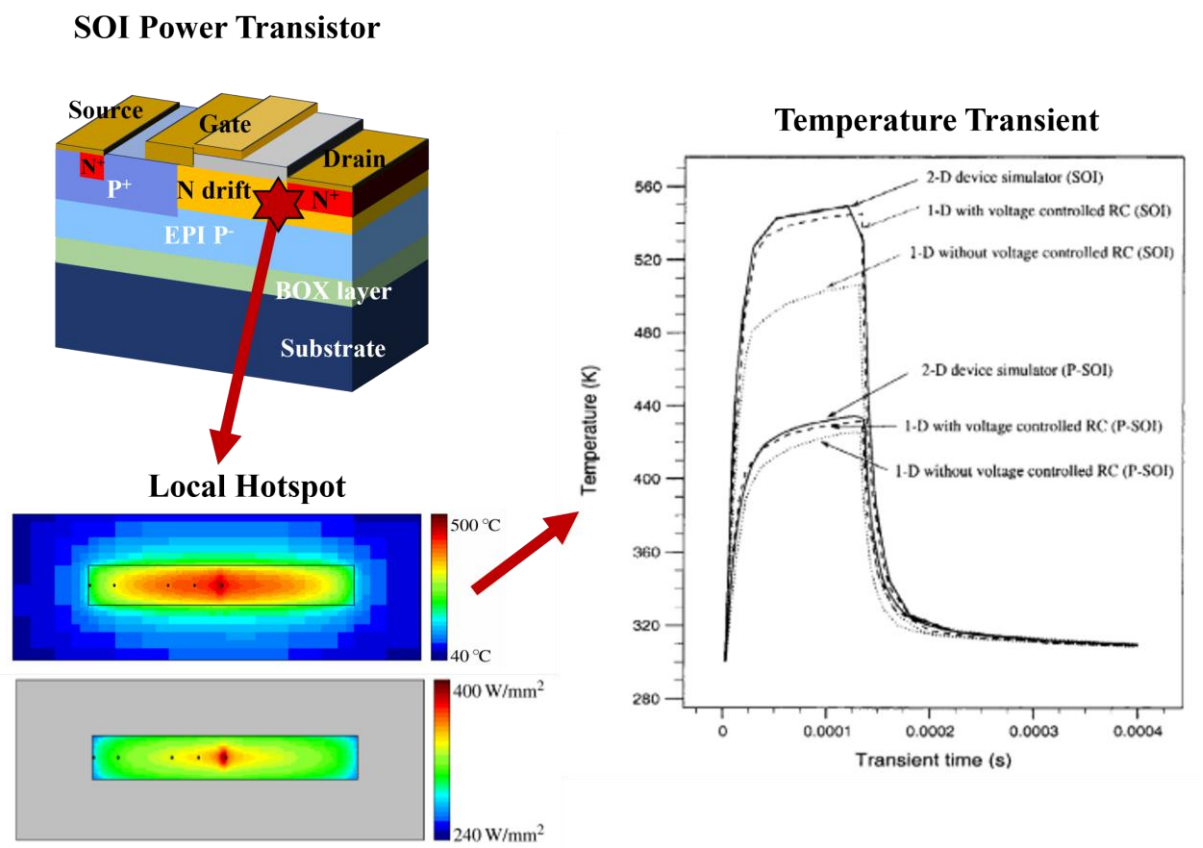

Figure 1-2. Self-heating effect in SOI power transistors. Bottom image adapted from [6]. *Right image adapted from* [5] *).*

field still exists in the channel region near the drain terminal, creating local "hotspots" with power density generally on the order of 0.1 kW/cm^2 to 1 kW/cm^2 [4], [5]. For certain applications with power MOSFETs (e.g., laterally diffused metal-oxide-semiconductor, LDMOS), the hotspot power density can reach 10 kW/cm^2 [6]. High power density in hotspot together with low thermal conductance components (e.g., buried oxide layer) can result in rapid temperature rise up to hundreds of kelvins within a few microseconds (as shown in Figure 1-2) [5]. Such high temperature can change the transistor electrical characteristics, such as threshold voltage, leakage current and breakdown voltage, which have a negative impact on device reliability. Besides, prolonged exposure to high temperature causes hot-carrier injection (HCI), which can lead to deterioration in transistor performance and lifetime [7], [8]. Furthermore, once the transistor temperature reaches

the critical point (can be around 340 – 520 °C [9]), the device becomes thermally unstable as drain current increases with temperature due to a positive temperature coefficient. This leads to irreversible thermal runaway, resulting in permanent device failure and thus determining the safe operating area (SOA) [9], [10].

1.3 Overview of Conventional Cooling Methods for Microelectronic Devices

Nowadays, the power density in high-performance microelectronic devices has achieved 100W/cm^2. Increased power density and high power dissipation call for the development of cooling technologies in microprocessors. For many years, forced air cooling has been the most popular option in ICs because of its good compatibility, high reliability and low costs [11]. Great efforts have been devoted to the design optimization of air-cooling components such as heat sink, thermal interface, heat spreader and fan. However, due to the nature of poor thermal properties of air, the cooling capabilities of air cooling cannot be significantly improved. Sain *et al.* predicted that air cooling has the limit of only 37 W/cm^2 in heat removal capacity when optimal geometries are employed [12]. Xu *et al.* showed that air cooling can only handle a maximum allowable CPU power of only 340 W although ideal situations, including air flow, power distribution, package materials and configurations, are considered [13].

Compared to air cooling, liquid cooling has demonstrated greater cooling potential [14]. Liquid coolants (e.g., water) have much higher thermal conductivity, density and heat capacity compared to air, which make them desirable to be efficient heat exchange media. Heat removal capability of up to 170 W/cm^2 can be achieved using the state-of-the-art liquid cooling design with the implementation of microchannels [15]. On the other hand, two-phase cooling, which takes

advantage of large latent heat during the liquid-vapor phase change process, shows additional advantages over liquid cooling and has recently become the research of interest [16]. Two-phase cooling to address a chip power of 2.9 kW and surface heat fluxes up to 910 W/cm^2 was successfully performed by Drummond *et al* [17]. However, both liquid and two-phase cooling methods have their own set of challenges, including complexity, risk of leaks and high maintenance cost.

In addition to the active cooling methods mentioned above, researchers have explored other passive cooling methods such as phase change materials (PCMs) [18] and thermal interface materials (TIMs) [19]. Those cooling methods, while promising, are primarily used in conjunction with active cooling methods, which cannot significantly enhance cooling performance.

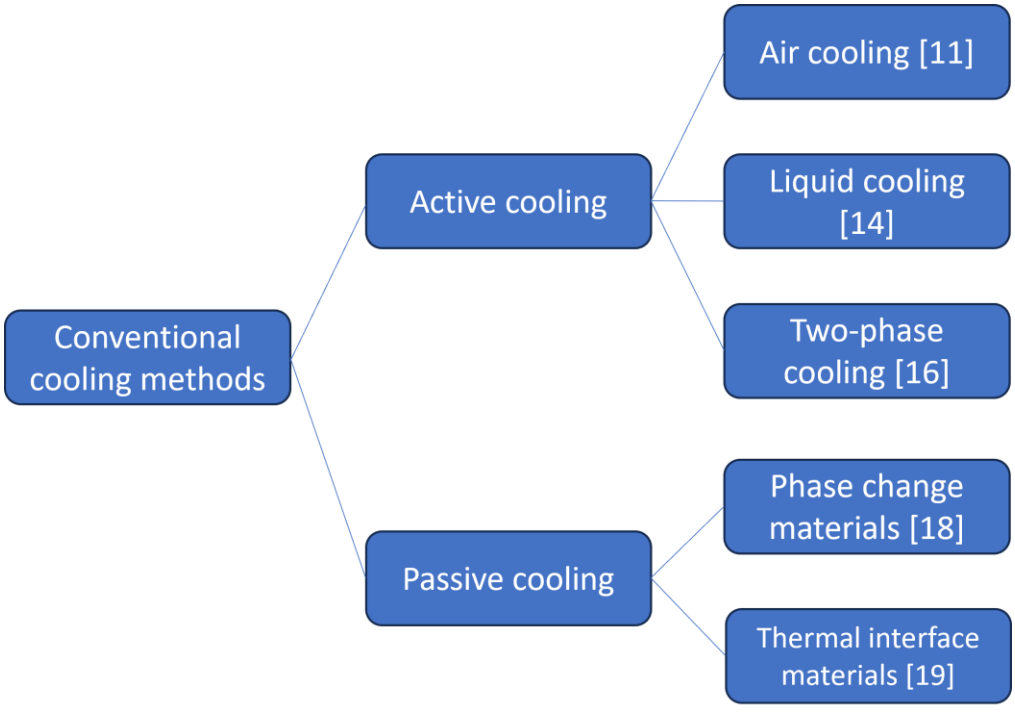

Figure 1-3. Category of conventional cooling methods .

1.4 Fundamentals of Thermoelectric Cooling

Compared to conventional cooling methods, thermoelectric cooling (i.e., Peltier cooling) is a novel active cooling method in microelectronic systems due to its high cooling power [20], solid-state operation [21] and scalability [22]. The principle behind the Peltier effect is that when electrons move from a material with lower energy level to a material with higher energy level, they absorb heat from the surroundings and induce cooling. As a result, Peltier cooling is an active cooling method happening near the metal-semiconductor junction, which can address the local hotspot issues in ICs.

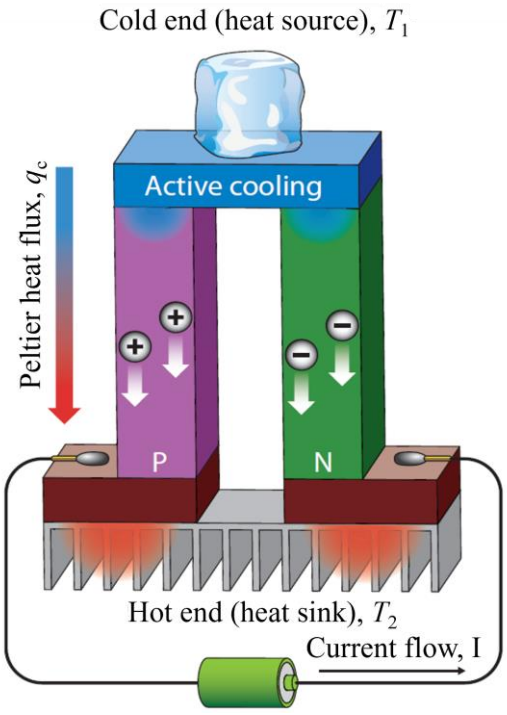

Figure 1-4. Schematic of a thermoelectric cooler system. Figure adapted from [74].

In steady state, the cooling power of a 1-D thermoelectric cooler (TEC) with both p-type and n-type legs (as shown in Figure 1-4) is given by [23]:

$$q_C = (S_p - S_n)IT_1 - K(T_2 - T_1) - \frac{I^2R}{2} \quad (1)$$

where q_C is steady-state cooling power, S_p is the Seebeck coefficient of the p-type leg, S_n is the Seebeck coefficient of the n-type leg, T_1 is absolute temperature at the cold end (heat source), T_2 is absolute temperature at the hot end (heat sink), I is TEC current, K is thermal conductance, and R is electrical resistance. During the Peltier cooling event, the active heat flux ($\propto I$) overcomes Fourier heat conduction and the concomitant Joule effect ($\propto I^2$). Therefore, a trade-off in I exists which results in a particular TEC current/voltage to provide the greatest cooling power, which can be expressed by [23]:

$$I_q = \frac{(\alpha_p - \alpha_n)T_1}{R} \tag{2}$$

$$(q_C)_{\mathrm{MAX}} = \frac{(\alpha_p - \alpha_n)^2 T_1^2}{2R} - K(T_2 - T_1) \tag{3}$$

In TECs with efficient heat sink, the cold end temperature, T_1, can be considered constant. Therefore, the temperature difference between hot end and cold end can be written as:

$$T_2 - T_1 = \frac{(\alpha_p - \alpha_n)^2 T_1^2}{2RK} - \frac{q_c}{K} \tag{4}$$

In practical applications, one should consider thermal contact resistance, R_{th}. Therefore, equation (4) should be modified as [24]–[26]:

$$T_2 - T_1 = \frac{(\alpha_p - \alpha_n)^2 T_1^2}{2RK} - q_c\left(\frac{1}{K} + R_{th}\right) - R_{th}IV - \frac{R}{2K}I^2 \tag{4*}$$

Apparently, the maximum temperature difference $(T_2 - T_1)_{\mathrm{MAX}}$ can be obtained only when $q_c = 0$. In literature, such $(T_2 - T_1)_{\mathrm{MAX}}$ is often used to evaluate the cooling performance of a specific TEC device.

While steady-state studies have contributed to optimizing the TEC performance, many studies have put their focus on the transient behavior of the Peltier effect [20], [27]–[30]. In contrast to steady-state Peltier cooling, transient Peltier cooling exhibits spatial and temporal mismatches between the Peltier effect and concomitant Joule effect. Specifically, Peltier cooling occurs at the cooler junction simultaneously with TEC activation, whereas volumetric Joule heating takes place throughout the TEC with a delayed response due to heat diffusion. Figure 1-5 demonstrates that by combining an optimal steady-state TEC current with a transient TEC current pulse, one can achieve more pronounced transient cooling despite a delayed temperature overshoot. This phenomenon is referred to as supercooling effect and can be utilized for addressing sudden power dissipation and transient thermal shock.

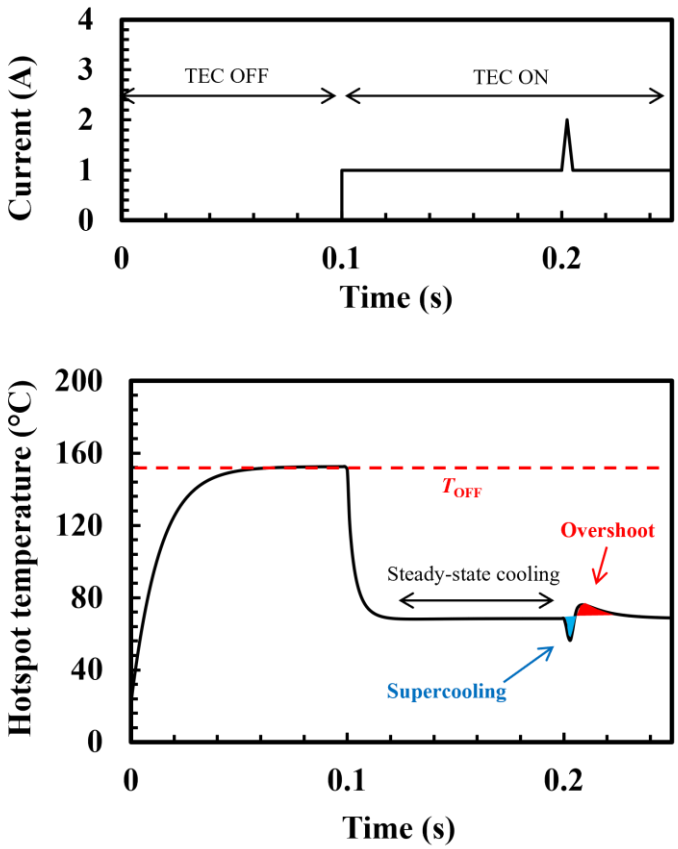

Figure 1-5. TEC supercooling effect

1.5 Development of Thermoelectric Coolers in Microelectronic Systems

Due to the local cooling effect, TECs are highly desirable for cooling specific hotspot regions (typically less than 500 μm × 500 μm) on a die. Over the past two decades, a lot of efforts have been made in developing thermoelectric coolers (TECs) suitable for microelectronic systems. Figure 1-6 depicts several featured TEC designs and the corresponding studies: back in 2005, Prof. Shakouri's group experimentally demonstrated a Si/SiGe superlattice TEC that has 6.9 K temperature reduction at 100°C [31]. In 2007, Prof. Bar-Cohen's team developed an analytical model for silicon-based on-chip thermoelectric microcoolers [32], which utilized the most common material in semiconductor industry, single crystal silicon, to address high hotspot heat

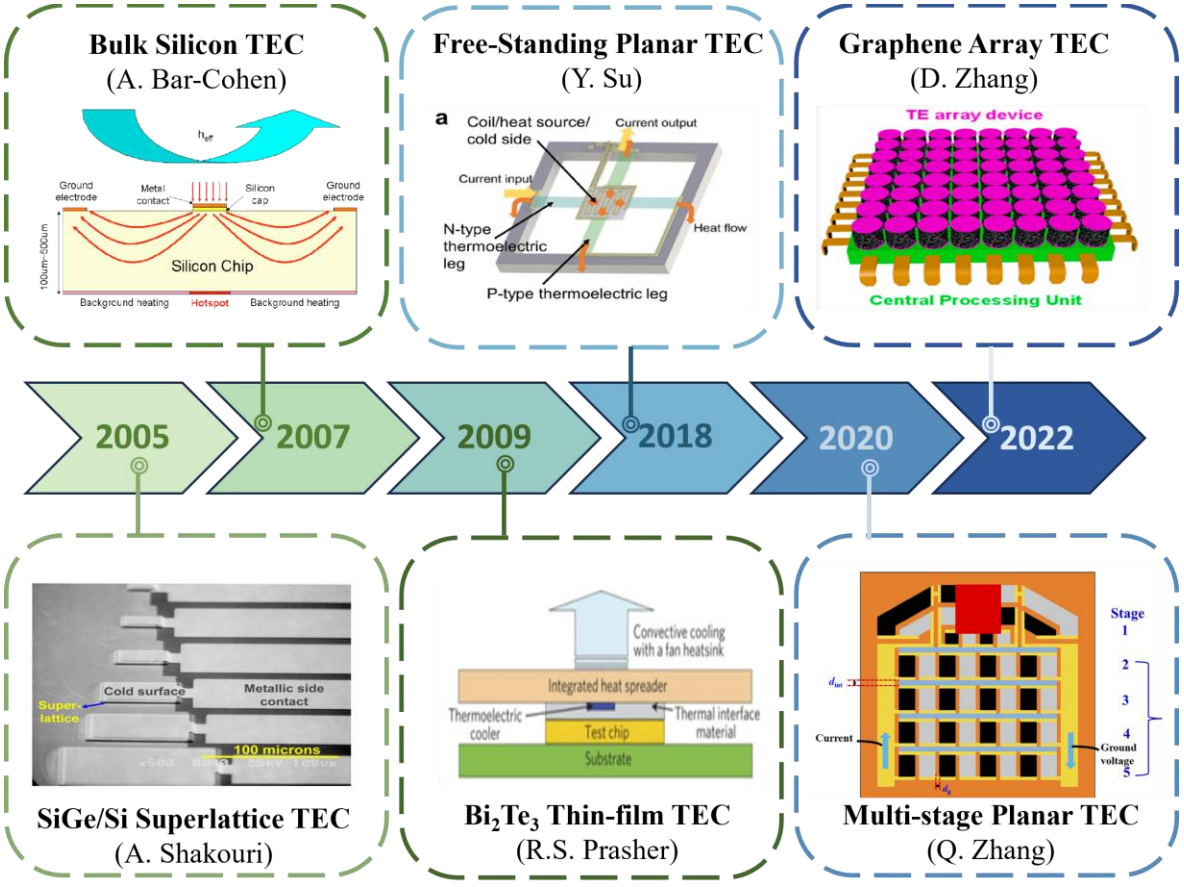

Figure 1-6. Development of thermoelectric coolers. Figures adapted from [31]–[37].

flux. In 2009, Prof. Prasher's group for the first time examined the cooling performance of a Bi_2Te_3 superlattice-based thin-film TEC integrated into a contemporary electronic package. 15 K cooling was observed at the hotspot region with a high heat flux of about 1300W/cm^2 [33]. In 2018, Su *et al.* demonstrated a free-standing SiGe-based planar thin-film TEC [34], [35]. In this paper, the SiGe thin films were growth by LPCVD on top of a sacrificial SiO_2 layer. After SiO_2 removal using RIE, suspended SiGe thin films were created, which can effectively prevent parasitic heat loss. In 2020, Nie *et al.* showed the numerical simulation and structural optimization of multi-stage planar TECs [36]. By optimizing the leg thickness, stage number as well as the p-n pair number, one can realize a maximum temperature reduction of 8.2K. Recently in 2022, Zhang et al. demonstrated a TEC array using graphene-based sponges as the thermoelectric material [37]. After mixing graphene with PDMS, those sponges possess high Seebeck coefficient, high elasticity as well as excellent mechanical stability.

Among those TECs, the material selection and device performance are two major considerations. Only those materials which combine high ZT and good compatibility with the conventional microfabrication process are desirable for thermoelectric microcoolers. On the other hand, the device performance can be enhanced by lowering thermal contact resistance and minimizing parasitic heat loss. Moreover, those TECs should allow easy fabrication using macro- and micro- fabrication processes.

1.7 Vertical TEC vs. Lateral (planar) TEC

Despite many ways to distinguish different TECs, one way is based on the orientation of the active heat flux. The vertical designs, which generates cooling flux in the vertical direction, are the most commonly applied in literature. Typical vertical TECs include thin-film thermoelectric coolers (TFTECs), which take advantage of nanostructured superlattices such as Bi_2Te_3/Sb_2Te_3, PbSeTe/PbTe, and SiGe/Si [24], [38], [39]. However, vertical TECs have several major drawbacks. Firstly, relatively large thermal contact resistance exists in vertical direction due to thin film deposition, leading to degradation of cooling performance. Secondly, due to the configuration constraints, the vertical TEC is placed on the opposite side of the substrate with a distance of 100 – 300 μm between the cooler and the cooling object [40], [41], resulting in substantial parasitic heat loss and insignificant cooling performance. This situation worsens for SOI devices as their BOX layer with low thermal conductivity is detrimental to vertical heat dissipation. Lastly, many transistor and transistor-based array systems experience dynamic change in heating conditions based on different working environments and operating conditions, yet many vertical TECs assume a fixed hotspot location and steady-state heating conditions, which cannot provide on-demand temperature control [42].

On the other hand, lateral designs redistribute heat laterally, which takes advantage of the sustained temperature gradient and the non-uniformity of power density in lateral direction. Besides, due to the absence of intermediate substrate, the lateral TEC minimizes the distance between the cooler and the cooling object, thus resulting in less parasitic loss and more significant cooling performance. Figure 1-9 provides three possible TEC designs based on different configurations and orientations of active heat flux.

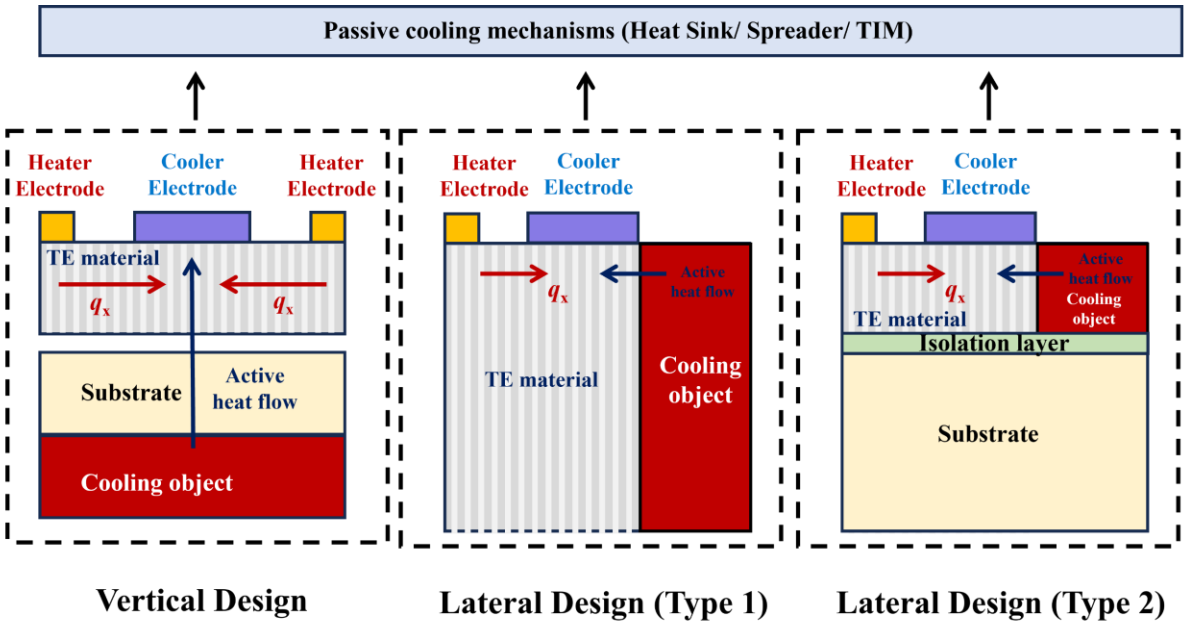

Figure 1-7. Holey silicon-based TECs in vertical and lateral configurations. The concept of vertical design is adapted from [51].

1.8 Holey Silicon as a Promising Thermoelectric Material

For thermoelectric materials, the thermoelectric efficiency can be defined by the thermoelectric figure of merit, ZT, which is given by:

$$ZT = \frac{S^2\sigma}{k}T \tag{5}$$

where S is the Seebeck coefficient, σ is electrical conductivity, k is thermal conductivity and T is absolute temperature. While great improvements in ZT has been achieved through nano-structuring [43] and controlled doping [44], most predominant thermoelectric materials (e.g., Bi_2Te_3) are based on complex materials [45], which are difficult to fabricate using conventional microfabrication processes and thus not suitable for microelectronic devices.

Holey silicon, on the other hand, is one of the few promising thermoelectric materials that offer both high thermoelectric efficiency (maximum theoretical ZT ~ 0.47 at room temperature [46]) and excellent compatibility with the bipolar-CMOS-DMOS (BCD) technology. As shown in Figure 1-8, the holey silicon is a nanostructured silicon with arrays of nanoscale vertical holes, which can be readily created using conventional lithography and deep reactive ion etching (DRIE) techniques. This special morphology of holey silicon effectively enhances lateral phonon scattering, leading to a significant reduction in in-plane thermal conductivity. Experimental results and analytical models [46]–[50] have demonstrated that the holey silicon thin film, possessing optimal neck size and porosity, can achieve low in-plane thermal conductivity of 1 – 10 W/m·K

at room temperature. Moreover, due to its relatively short electron mean free path, holey silicon can retain high electrical conductivity and Seebeck coefficient from bulk silicon.

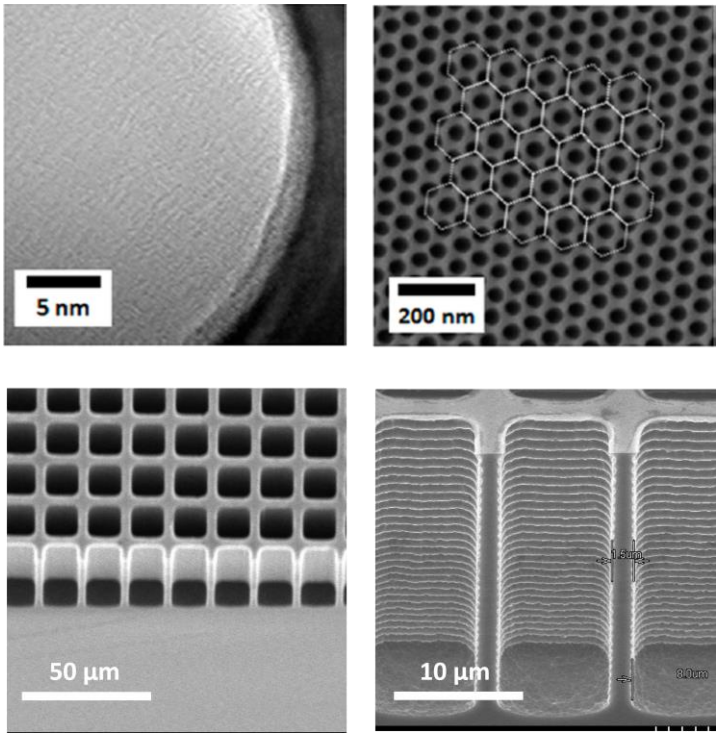

Figure 1-8. Holey silicon nanostructures with (top) 20nm neck size and 200 nm thickness and (bottom) 2µm neck size and 30 µm thickness. Top two figures adapted from [48].

1.9 Research Goals and Objectives

The goal of this thesis is to develop a holey silicon-based lateral thermoelectric cooler system that provides efficient dynamic cooling for transistors. We choose the planar power transistor (i.e., LDMOS) as our cooling object due to its high power dissipation and the vulnerability to thermal failure (i.e., thermal runaway). By optimizing the TEC design, we try to extend the allowable power of the power transistor. Steady-state and transient simulations using AC/DC and heat transfer modules in COMSOL Multiphysics are performed to evaluate and optimize the cooling performance of the transistor-TEC system.

The following objectives are achieved to fulfill the goals:

1. Design and modelling of the holey silicon-based lateral TEC and its transistor-TEC system
2. Steady-state study of the transistor-TEC system, with the optimization in geometrical parameters, material properties and boundary conditions.
3. Transient study of the transistor-TEC system, with the investigation in different current amplitudes, locations, durations, and pulse shapes.
4. Preliminary study of the transistor-TEC array

Chapter 2

Basic Model Setup

2.1 Device Configuration

2.1.1 Transistor configurations

When designing a TEC for microelectronic systems, the first step is to define the cooling object. Our research of interest lies in the planar power transistors (i.e., laterally diffused metal-oxide-semiconductors, LDMOSs), as they dissipate a large amount of heat in a wide range of applications such as 5G wireless infrastructure, satellite communications and radar systems. Nevertheless, even a specific transistor type like an LDMOS can have thousands of configurations and dimensions based on the actual applications, which makes it difficult to develop a general TEC design. To address this issue, we weaken the concept of the LDMOS configuration and only consider the cooling object as a p-type single crystal silicon block without detailed components such as gate, source and drain. Following this, we extract the essential hotspot features (e.g., location and dimensions) in an LDMOS transistor as they are crucial in heat transfer analysis. Generally, the hotspot is located beneath the gate electrode near the drain terminal within the channel region of an LDMOS transistor. For simplicity, we assume that this hotspot is positioned at the center of the transistor in xy-plane and a few micrometers below its top surface. Figure 2-1 depicts the transformation from a specific LDMOS into a generalized model implemented in this

study. In the following chapters, more detailed information of the geometrical dimensions will be discussed.

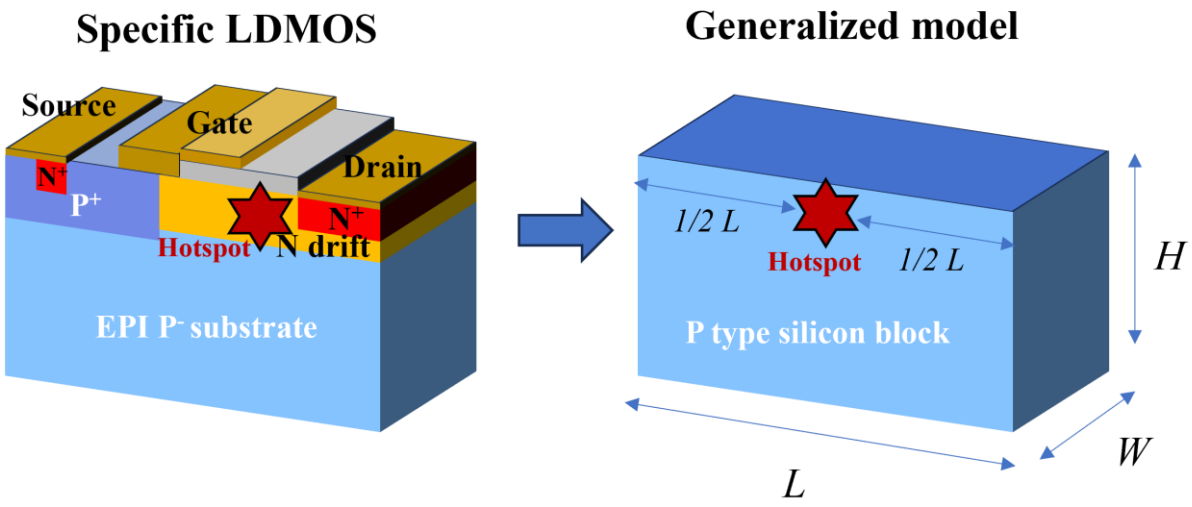

Figure 2-1. Transformation from a specific LDMOS device to a generalized model.

2.1.2 TEC configurations

Previous vertical TEC designs have encountered several issues such as large contact resistance and insufficient cooling. In contrast, a lateral TEC design may show greater promise in cooling microelectronic devices. Fabricated by conventional lithography and deep reactive ion etching (DRIE) techniques using silicon wafers, holey silicon contains vertical nanoscale holes, resulting in enhanced lateral phonon scattering and reduced in-plane thermal conductivity. The special morphology of holey silicon makes it desirable to become a lateral TEC. Figure 2-2 shows the schematic of a transistor-TEC system using the holey silicon-based lateral TEC. Only one half of the device is shown due to symmetry. The holey silicon-based lateral TEC comprises three components: the Peltier heater, the p-type holey silicon region and the Peltier cooler. Such horizontal alignment of the TEC allows for direct contact between the cooler and the transistor, enabling more efficient heat removal compared to the conventional vertical TEC designs.

When the TEC is ON, positive current goes through the p-type holey silicon region from the Peltier cooler, resulting in temperature reduction at the metal-semiconductor junction and its adjacent LDMOS. Such current is then terminated at the Peltier heater, which is situated far away from the LDMOS, leading to a simultaneous temperature rise. During the Peltier effect, the active heat flow ($\propto I$) shares the same direction with the TEC current (from right to left), while the Fourier heat flow is in the opposite direction (from left to right).

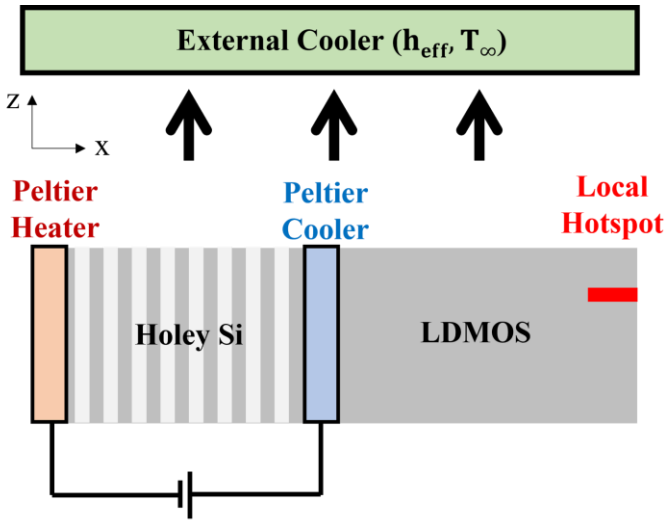

Figure 2-2. Schematic of a transistor-TEC system using lateral holey silicon-based TEC (only one half of the device is shown due to symmetry).

2.2 Material Selection and Their Properties

2.2.1 Thermal conductivity of holey silicon

Compared to bulk silicon, holey silicon exhibits a reduction in thermal conductivity due to the enhancement of phonon scattering. Furthermore, it demonstrates anisotropy between cross-plane and in-plane thermal conductivity, which is supported by both experimental investigations and theoretical analyses. According to the spectral scaling model, the cross-plane thermal

conductivity in holey silicon is influenced by the neck size (n) and holey silicon thickness (t), which is given by [51]:

$$k_z = \int_0^{\omega_D} k_{z,\infty}(\omega, n) \times \left(1 + \frac{\lambda_{z,\infty}(\omega, n)}{0.5t}\right)^{-1} d\omega \tag{6}$$

where ω stands for the phonon frequency, ω_D is the Debye cut-off frequency, $k_{z,\infty}(\omega, n)$ and $\lambda_{z,\infty}(\omega, n)$ represents the thermal conductivity and the phonon MFP of the corresponding silicon nanowire. The last two variables can be further derived by the following Landauer formalism [52]:

$$k_{z,\infty} = \frac{2L}{\pi^2 d_{eff}^2} \int_0^{\infty} \left(\frac{N_1(\omega)}{1 + \frac{L}{\lambda_1(\omega)}} + \frac{N_2(\omega)}{1 + \frac{L}{d_{eff}}}\right) \frac{h^2(\omega)}{k_B T^2} \times \frac{\exp(\hbar\omega \,/\, k_B T)}{(\exp(\hbar\omega \,/\, k_B T) - 1)^2} d\omega \tag{7}$$

where T denotes the temperature, L is the length and d_{eff} stands for the characteristic diameter that represents an equivalent cross-sectional area of a nanowire. For holey silicon, $d_{eff} = \sqrt{\frac{\sqrt{3}}{\pi} p^2 - \frac{(p-n)^2}{2}}$, where p is the pitch size and n is the neck size. $\lambda_1(\omega)$ is the frequency dependent MFP which can be written as $\lambda_1(\omega) = 4\pi^2 \times \left(4B \frac{\hbar^2}{d_{eff}^3} \left(\frac{\omega}{\omega_D}\right)^2 + AB \frac{\hbar^2}{a_l^2 d_{eff}} \left(\frac{\omega}{\omega_D}\right)^4\right)^{-1}$, where a_l is the lattice spacing, A and B represent the dimensionless fitting parameters. Finally, the $\lambda_{z,\infty}(\omega, n)$ can be obtained by combining the scattering coefficients including lateral boundary scattering ($\lambda_B(\omega)$), Umklapp scattering ($\lambda_U(\omega)$) and point-defect scattering ($\lambda_D(\omega)$):

$$\lambda_{z,\infty}(\omega, n) = (\lambda_B(\omega)^{-1} + \lambda_U(\omega)^{-1} + \lambda_D(\omega)^{-1})^{-1} \tag{8}$$

Likewise, the in-plane thermal conductivity of holey silicon can be predicted using the spectral scaling model, given by [51]:

$$k_x = \int_0^{\omega_D} k_{x,\infty}(\omega, n) \times \left(1 + \frac{\lambda_{x,\infty}(\omega, t)}{0.5n}\right)^{-1} d\omega \tag{9}$$

where $k_{x,\infty}(\omega, n)$ is the in-plane thermal conductivity of the silicon thin film, which can be obtained by the Holland model [53]:

$$k_{x,\infty} = \frac{1}{3}\sum_j v_j^2 \int_0^{\frac{\theta_j}{T}} C_{V,j}(x_\omega, T)\left[\tau_j(x_\omega, T, t) \times F(\delta)\right] dx_\omega \tag{10}$$

The subscript j denotes the transverse and longitudinal phonon modes, v_j is the group velocity, $x_\omega = h\omega/k_B T$ is the non-dimensional phonon frequency. $C_{V,j}$ is the phonon specific heat per unit volume, θ_j stands for the Debye temperature and τ_j denotes the phonon relaxation time without the vertical phonon boundary scattering. $F(\delta)$ is the reduction function explained in [54]–[56].

Figure 2-3 shows the cross-plane and in-plane thermal conductivity of holey silicon derived by the spectral scaling model. A decrease in thickness or neck size results in a corresponding decrease in cross-plane or in-plane thermal conductivity. Specifically, when the holey silicon neck size is below 100 nm, the in-plane thermal conductivity can be as small as 1-10 W/m·K.

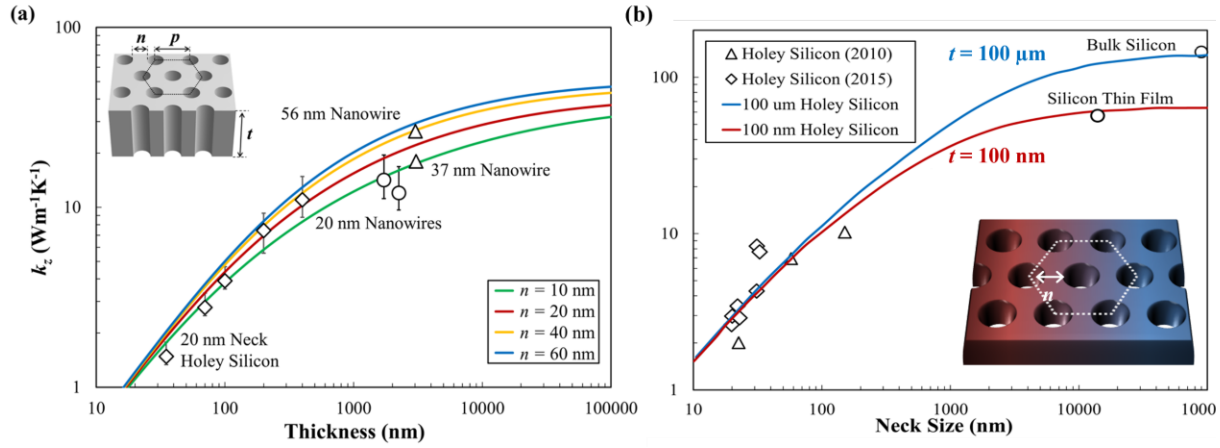

Figure 2-3. The (a)cross-plane and (b) in-plane thermal conductivity of holey silicon calculated by the spectral scaling model. Figures adapted from [51]

In this study, we utilize the pre-defined MATLAB program using the spectral scaling model to predict the anisotropic thermal conductivity of holey silicon. Two sets of constant thermal conductivity for holey silicon are employed: {k_x, k_y, k_z} = {2.6, 2.6, 35} W/m·K and {1, 1, 13} W/m·K, taking into account the parameters {neck size, thickness, porosity, temperature} = {20 nm, 30 μm, 30%, 300 K} and {10 nm, 30 μm, 30%, 300 K}, respectively. It should be noted that these calculated values are subject to certain assumptions and may vary due to other factors. Furthermore, the practical implementation of holey silicon dimensions should also consider feasibility requirements.

2.2.2 Electrical conductivity and Seebeck coefficient of holey silicon

The phonon mean free path (MFP) in silicon typically ranges between 200-300 μm [57], [58], whereas the electron MFP is significantly smaller at approximately 10 nm [59]. This low electron MFP in silicon enables the preservation of excellent electrical characteristics even when nano-structuring is employed.

In silicon, the electrical conductivity and Seebeck coefficient are strongly influenced by the doping concentration. While the electrical conductivity exhibits a positive correlation with the doping level, the Seebeck coefficient demonstrates an inverse relationship with the doping concentration due to the difference in Fermi level. Since the thermoelectric figure of merit, $ZT = S^2\sigma T/k$, involves the product of the Seebeck coefficient and electrical conductivity, it reaches its maximum value at a specific doping concentration.

Based on the Boltzmann transport theory, the electrical conductivity and Seebeck coefficient can be expressed by [60]:

$$\sigma = q_0^2 \int_{E_0}^{\infty} \left(-\frac{\partial f_0}{\partial E}\right) N(E) v(E) \lambda_0 \left(\frac{E}{k_B T}\right)^r dE \tag{11}$$

$$S = \frac{q_0 k_B}{\sigma} \int_{E_0}^{\infty} \left(-\frac{\partial f_0}{\partial E}\right) N(E) v(E) \lambda_0 \left(\frac{E}{k_B T}\right)^r \left(\frac{E - E_f}{k_B T}\right) dE \tag{12}$$

where q_0 is the elementary positive charge, k_B is the Boltzmann constant, λ_0 is the electron MFP, $v(E)$ is the electron group velocity, E_f is the Fermi level, T is the absolute temperature, f_0 is the Fermi-Dirac distribution function. For p-type semiconductor, $N(E)$ stands for the density of states of the conductance band; for n-type semiconductor, $N(E)$ represents the density of states of the valence band.

The one-dimensional density of state can be derived using the following equation:

$$N(E) = \frac{4\pi(2m_n^*)^{\frac{3}{2}}}{h^3}\sqrt{E - E_c} = \frac{4\pi[2(1.08)(9.11 \times 10^{-31})]^{\frac{3}{2}}}{(6.625 \times 10^{-34})^3}\sqrt{E - E_c} \tag{13}$$

where m_n^* is the effective mass for electrons, m_n^*=1.08m_0, m_0= 9.11×10^{-31}, h is plank constant, h= 6.625×10^{-34}, E_c is the conductance band edge. For the three-dimensional density of state, the density functional theory (DFT) can be used to calculate the silicon band structure which then provides the density of state. With this method, the group velocity, $v(E)$, can be derived at the same time.

For non-intrinsic semiconductor, the Fermi level, E_f, can be express as:

$$E_C - E_F = k_B T ln(\frac{N_C}{n_0}) \tag{14}$$

where N_C is the effective conductance density of states and n_0 is the electron carrier concentration. In N-type semiconductor, $N_d >> n_i$, where n_i is intrinsic carrier concentration (~10^{10}/cm^3) and N_d

is donor concentration. Therefore, $n_0 \approx N_d$, and the equation above can be further approximated as:

$$E_C - E_F = k_B T ln\left(\frac{N_C}{N_d}\right) \xrightarrow{N_C = 2.8\times10^{19}\ cm^{-3}\ @\ 300K} E_C - E_F = 0.0259 ln\ (\frac{2.8 \times 10^{19}}{N_d}) \quad (14^*)$$

The electron relaxation time is the reciprocal of the electron scattering, which includes electron-phonon scattering ($\Lambda_{e\text{-}p}$), ionic impurities (Λ_{II}), intervalley scattering (Λ_{IV}), boundary scattering (Λ_{BS}), optical phonon (Λ_{opt}), and plasma scattering (Λ_{PS}). Assuming that $\Lambda_{e\text{-}p}$, Λ_{II}, Λ_{BS} are the dominant scattering mechanisms in holey silicon, based on Matthiessen's Rule, it has:

$$\Lambda(E) = \Lambda_{e-p} + \Lambda_{II} + \Lambda_{BS} \quad (15)$$

$$\tau(E) = \Lambda(E)^{-1} \quad (16)$$

Figure 2-4 shows the electron transportation calculation of bulk silicon using ab initio theory. We use the open-source computational model 'ThermoElectric.py' in GitHub [61]. Figure 2-4 (a) shows that the silicon electron MFP can range from 2 to 10 nm under different doping levels, which is consistent with the previous experimental data [59]. Figure 2-4 (b) shows the positive correlation between the electrical conductivity and the doping concentration, where the experimental data in [49] is also compared. Figure 2-4 (c) illustrates the negative correlation between the Seebeck coefficient and the doping concentration. A comparison among the theoretical data with and without the effective medium theory (EMT), as well as the experimental data in [46], [49] is shown. Eventually, in Figure 2-4 (d), a maximum power factor is shown when doping concentrations is between 1×10^{19} and 1×10^{20} cm^{-3}, a consistent result for both our model and the previous model in [46].

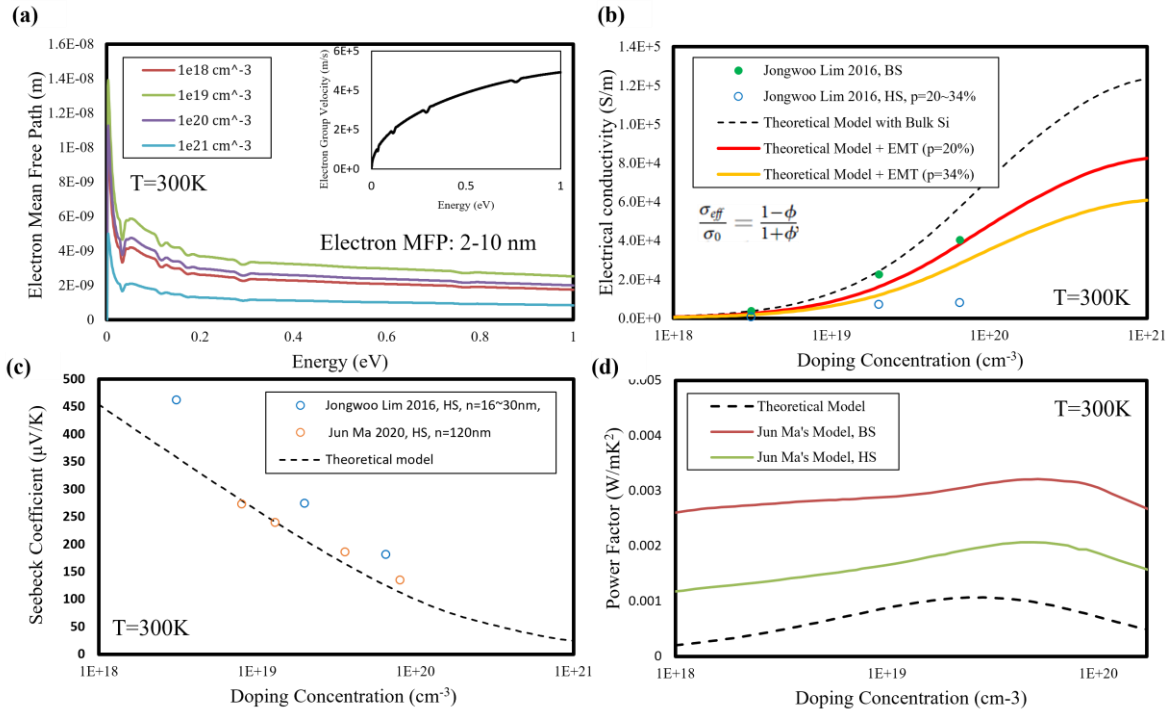

Figure 2-4. Electric transport computation of bulk silicon using ab initio theory. (The computational model and its python code adapted from [61] *)*

In this study, we utilize the documented electrical properties [62]–[64] of p-type silicon with a doping concentration of 2.5×10^{19} cm^{-3}, yielding an electrical conductivity of 2.52×10^{4} $S\cdot m^{-1}$ and a Seebeck coefficient of $440\mu V\cdot K^{-1}$. Finally, according to the effective medium theory, we derive an effective electrical conductivity of 1.36×10^{4} $S\cdot m^{-1}$ with a porosity level of 30%.

2.2.3 Other material selection and their properties

The LDMOS region is assumed to be a p-type single crystal silicon block, as previously discussed. While the electron transport properties in silicon (i.e., the Seebeck coefficient and electrical conductivity) remain unchanged as holey silicon, the phonon transport properties (i.e., the cross-plane and in-plane thermal conductivity) are slightly different. Detailed analytical model using the Boltzmann transport equation has been explained in [65], [66]. In this study, we

assign a constant thermal conductivity of 65 - 130 W/m·K to the silicon, taking into account the transistor thickness on the order of 10 μm and a temperature range of 300-600 K.

The Peltier cooler and Peltier heater are the thermoelectric electrodes made of metals, which typically require low absolute Seebeck coefficient in order to maximize the thermoelectric performance. Although we understand that different metals may possess various electrical and thermal properties and can further influence the TEC performance, we believe that this effect is subtle. Therefore, we choose aluminum or copper as the electrode material, given the fact that they are two of the common metals in microfabrication. We use the pre-defined temperature-dependent material properties in the COMSOL material library: for aluminum, the thermal conductivity, specific heat and electrical conductivity at 300K are 237 W/m·K, 904 J/kg·K and 3.6×10^7 S/m, respectively. For copper, those values are 385 W/m·K, 384 J/kg·K and 5.8×10^7 S/m, respectively.

2.3 Other Important Conditions

2.3.1 Heating conditions (transient study only)

To make reasonable analysis, the physical setting in the model (e.g., the initial & ambient temperature, the amplitude of heat pulse) should reflect the actual self-heating behavior of the transistor. Here we demonstrate three possible scenarios of the self-heating event:

Single heat pulse with local background temperature (Figure 2-5): In this scenario, the chip and its environment have achieved a steady-state high temperature. At this time, a single heat pulse with high power density (e.g., 50kW/cm^2) is applied to the system, which may reach the critical temperature and result in thermal failure (i.e., thermal runaway). In this case, one should set the ambient temperature and the initial temperature to a high value (e.g., 100 - 300°C) in order

to mimic the extreme background temperature condition. This scenario can be applied to the high-power cellular station or satellite infrastructure in extreme environments.

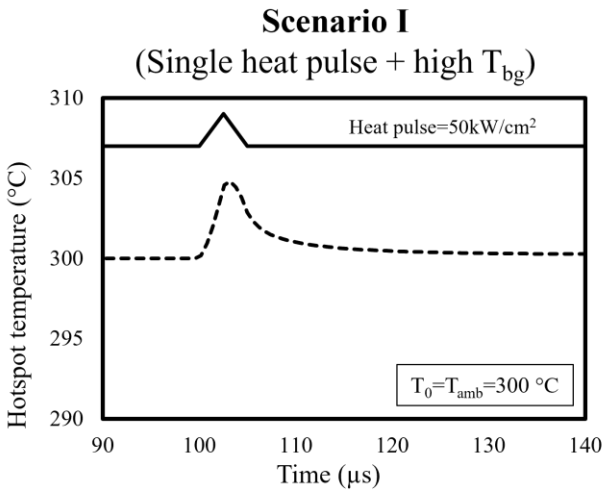

Figure 2-5. Single heat pulse with local background temperature

Consecutive heat pulse starting with room temperature (Figure 2-6): In this scenario, the transistor undergoes multiple switching which generates consecutive heat pulses. Every single heat pulse dissipates a certain amount of heat (e.g., 50kW/cm^2). Because of the large duty cycle and insufficient cooling, the hotspot temperature keeps increasing and finally results in thermal failure (i.e., thermal runaway). In this case, one should set the ambient and initial temperature to

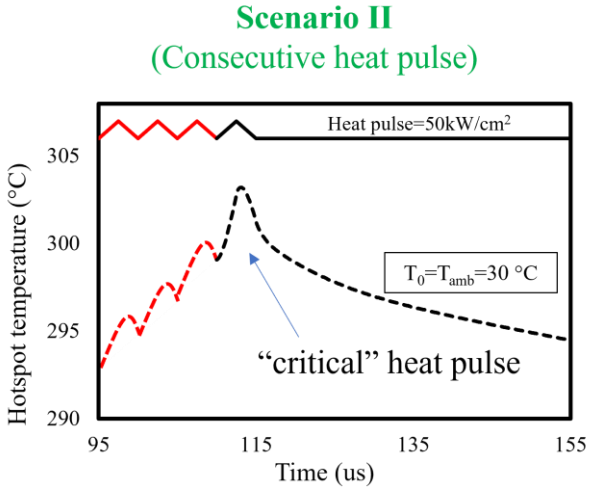

Figure 2-6. Consecutive heat pulse starting with room temperature.

room temperature (e.g., 30 °C) to mimic the initial situation without device activation. This scenario can happen because of system overspeed or cooling-system failure in signal amplification systems.

Single heat pulse with local constant heat flux (Figure 2-7): In this scenario, the chip has reached a steady-state high temperature where a constant heat flux (e.g., 30kW/cm^2) is dissipated. At this moment, a single heat pulse with moderate power density (e.g., 10kW/cm^2) is applied to the system, which may exceed the critical temperature and result in thermal failure (i.e., thermal runaway). In this case, one should set the ambient and initial temperature to room temperature (e.g., 30 °C). Besides, a constant heat flux should be applied at the hotspot. This scenario can occur when the power device is subject to sudden overload or overspeed.

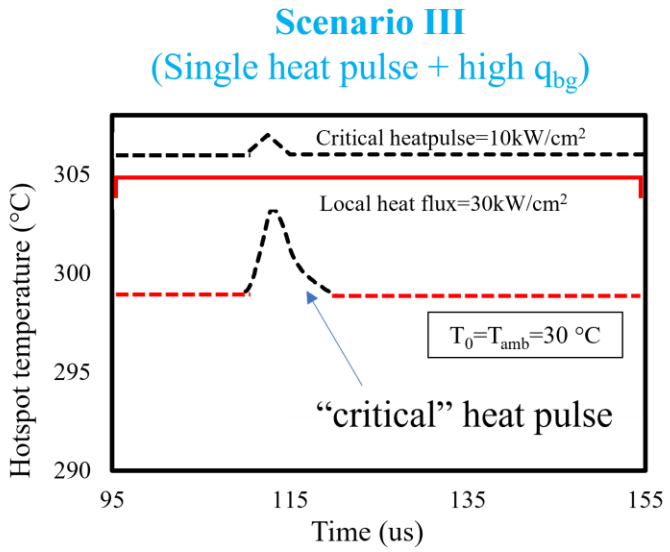

Figure 2-7. Single heat pulse with local constant heat flux

2.3.2 Boundary conditions

The determination of the boundary conditions in our study is based on the actual environment of the holey silicon-based lateral TEC. Typically, an efficient heat sink is connected to the top surface of the transistor layer, while a thermally insulated passivation layer covers the

opposite surface. Taking these factors into consideration, we assume a convective boundary for the top boundary condition and set an insulated boundary for the bottom.

The boundary condition of the side may undergo different situations: for a transistor-TEC matrix with multiple transistor-TEC units, the side may exhibit highly symmetric behavior in electrical and thermal transport. Therefore, an insulated boundary is more appropriate; for a transistor-TEC system situated next to high thermal impedance components, such as shallow trench isolation (STI) or air gap, an insulated boundary is also recommended. Moreover, to simulate a symmetric structure accurately, an insulated boundary is necessary. However, for a transistor-TEC system without any heating or insulation object on the side, a laterally conductive boundary is preferable. In this case, one should create a dummy silicon die to enclose the transistor-TEC system in order to provide perfect lateral thermal and electrical conduction. **Unless otherwise specified, we apply the insulated boundary on the side.**

More specific information regarding the boundary conditions, including the ambient temperature and the convective heat transfer coefficient, will be discussed in the upcoming sections.

Chapter 3

Steady-state TEC Cooling

While the actual interplay between the holey silicon-based lateral TEC and the power transistor may involve dynamic changes in thermal and electrical properties, a steady-state study can provide a preliminary and time-efficient evaluation of the transistor-TEC system. This ensures the parametric study in geometry and paves the way to further transient optimization.

In steady-state simulation, the heating conditions, TEC conditions and other boundary conditions are considered time independent. The following governing equations depict the physical assumptions and correlations within the coupled thermal-electrical model:

$$\nabla \cdot \boldsymbol{q} = Q_e \tag{17}$$

$$Q_e = \boldsymbol{J} \cdot \boldsymbol{E} \tag{18}$$

where $\boldsymbol{q}$ is the heat flux vector ($\nabla \boldsymbol{q}$ is the heat flux divergence) , Q_e is the internal heat source, which is the dot product of current density vector ($\boldsymbol{J}$) and electric field strength vector ($\boldsymbol{E}$).

The above equations apply to general thermal-electrical models. Specifically, when the thermoelectric effect is considered, the heat flux vector and current density vector should be further expressed as:

$$\boldsymbol{q} = ST\boldsymbol{J} - k\nabla T \tag{19}$$

$$\boldsymbol{J} = -\sigma(\nabla V + S\nabla T) \tag{20}$$

where S is the Seebeck coefficient (S), T is the temperature (∇T is the temperature gradient), k is the thermal conductivity, σ is the electrical conductivity and (T) and V is the applied voltage (∇V is the voltage gradient). As a result, the heat flux is coupled with the Peltier effect and the Fourier effect while the current density is coupled with the Seebeck effect and the Joule effect.

3.1 2D Simulation (Parametric Study)

3.1.1 Model selection

In the 2D simulations, the following two models are considered:

Boundary hotspot model: Figure 3-1 demonstrates the 2D transistor-TEC models implementing a boundary (line) hotspot. This model disregards the geometrical details of hotspot and simply treats the hotspot as a boundary. This model assumes that the hotspot is located far enough away from the TEC, where the size effect of the hotspot is minimal. Using this model with a boundary hotspot makes it accessible to the development of the corresponding 1D analytical model. More details regarding the corresponding 1D analytical model will be discussed in the upcoming sections.

Area hotspot model (selected for parametric study): Figure 3-2 demonstrates the 2D transistor-TEC models implementing an area hotspot. The geometrical information of the hotspot is included in this model. Specifically, the hotspot is positioned with a distance below the top surface (described by the variable "f"). And the hotspot dimensions is assumed to be 1 μm × 1 μm. Despite complexity, using this model with an area hotspot makes the result more reliable and is open for additional hotspot optimization.

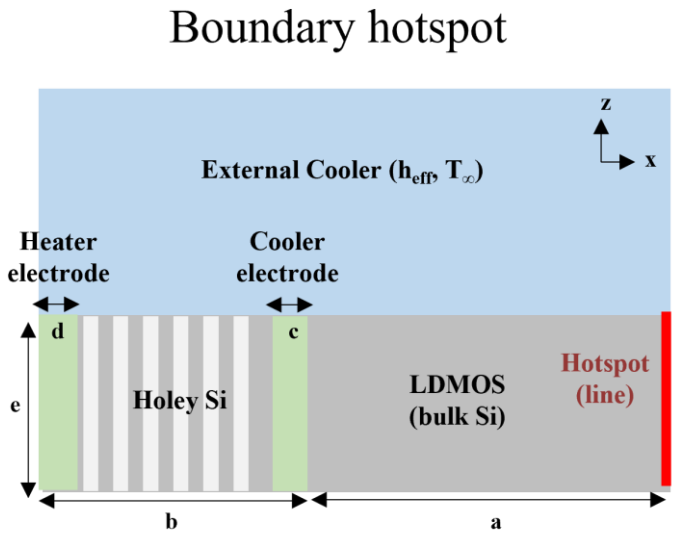

Figure 3-1. Schematic of 2D LDMOS-TEC system using boundary (line) hotspot.

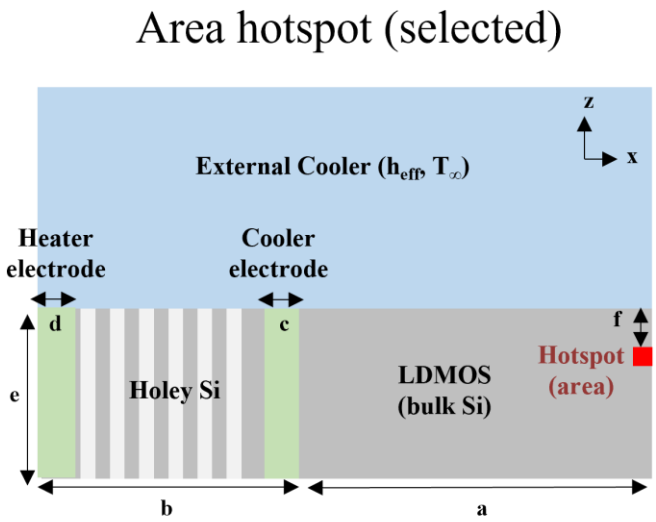

Figure 3-2. Schematic of 2D LDMOS-TEC system using area hotspot.

3.1.2 Reference simulation

The primary objective of the reference simulation is to demonstrate the fundamental functionality of the transistor-TEC system and provide a reference for further optimization. Table 3-1 presents the specific parameters utilized in this simulation:

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	20 µm
b	TEC length	80 µm
c	Length (x) of cooler	1µm
d	Length (x) of ground	1µm
e	Thickness (z) of chip	10µm
f	Distance between hotspot and top surface	1µm
L_{HS}	Length (x) of hotspot	1µm
T_{HS}	Thickness (z) of hotspot	1µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	1000 W/m^{2}K
Q	Hotspot heat flux	100 W/cm^2
T_∞	Ambient temperature	50 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	2.6 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	35 W/mK
k_{Si}	Thermal conductivity of bulk silicon	130 W/mK
k_{Cu}	Thermal conductivity of copper	385 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-1. Specific parameters used for 2D reference simulation.

A single parameter-sweep in TEC current is performed with the above parameters unchanged. As a result, the hotspot temperature is only influenced by the TEC current. As shown in Figure 3-3, the optimal current is 12 µA where a temperature difference of 1.02 °C is shown. For a current larger than the optimal current, the concomitant Joule heating surpass the Peltier cooling, leading to an undesirable temperature overshot.

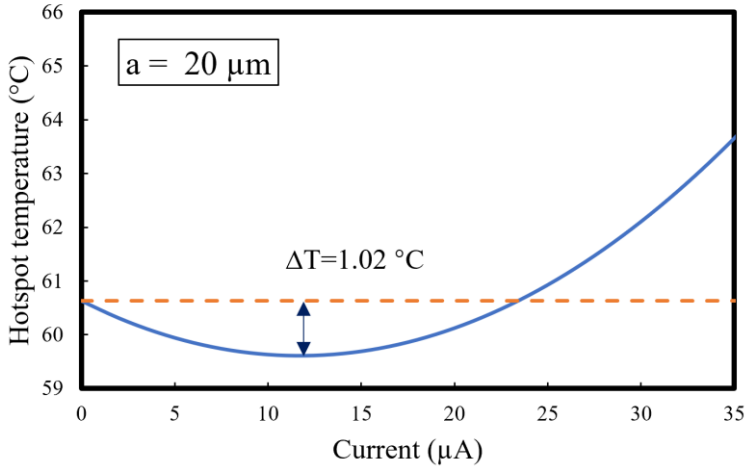

Figure 3-3. Relationship between hotspot temperature and TEC current.

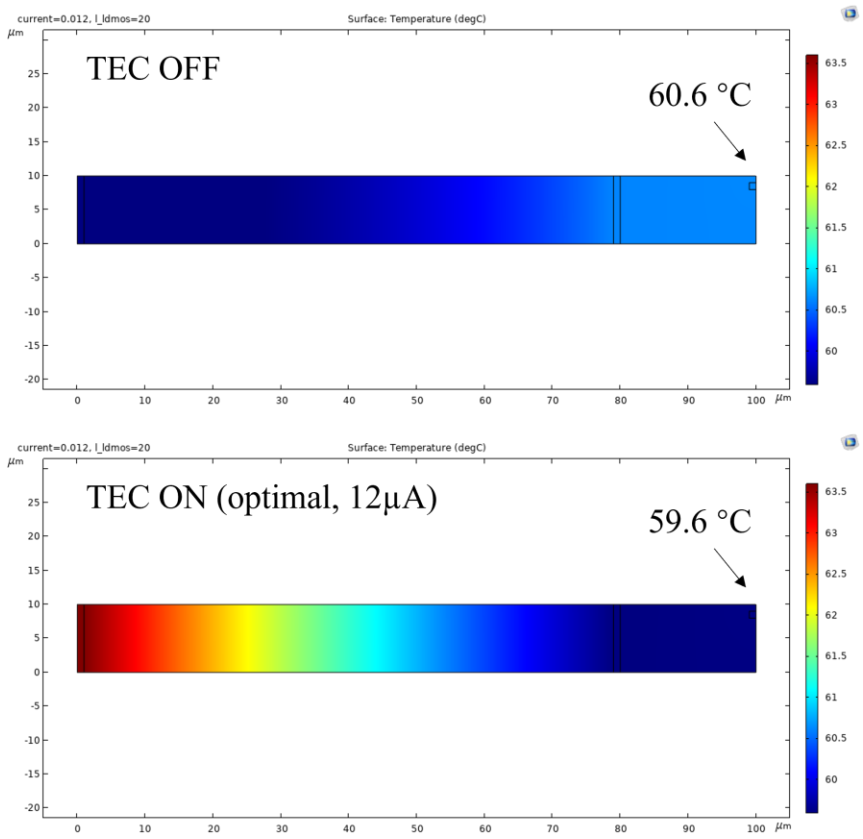

Figure 3-4. Comparison of temperature distribution between TEC OFF and TEC ON with optimal current.

The temperature distribution map is shown in Figure 3-4, where the TEC OFF and TEC ON situations are compared. In TEC OFF, the hotspot is the location where the highest temperature exists, which follows the Fourier heat conduction. However, in TEC ON, the TEC redistributes heat laterally, resulting in lower temperature in the cooler electrode and its adjacent hotspot region and higher temperature in the heater electrode.

3.1.3 Effect of hotspot distance

The hotspot distance (a) is a crucial parameter as it not only indicates the size of the transistor, but also implies the impact of the TEC on the hotspot. To further analyze the effect of

the hotspot distance, we perform a double parameter-sweep regarding the hotspot distance (a) and the TEC current. Table 3-2 presents the specific parameters utilized in this simulation:

Geometrical parameters

Parameters	Name	Value
a	**Hotspot distance**	**Varies 20-160 µm**
b	TEC length	80 µm
c	Length (x) of cooler	1µm
d	Length (x) of ground	1µm
e	Thickness (z) of chip	10µm
f	Distance between hotspot and top surface	1µm
L_{HS}	Length (x) of hotspot	1µm
T_{HS}	Thickness (z) of hotspot	1µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	1000 W/m^{2}K
Q	Hotspot heat flux	100 W/cm^2
T_∞	Ambient temperature	50 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	2.6 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	35 W/mK
k_{Si}	Thermal conductivity of bulk silicon	130 W/mK
k_{Cu}	Thermal conductivity of copper	385 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-2. Specific parameters used for hotspot distance study.

With optimal TEC current applied, a smaller hotspot distance shows greater temperature difference, as is shown in Figure 3-5. This suggests that the proximity between the cooler electrode and the hot spot plays a crucial role in maximizing the TEC cooling performance.

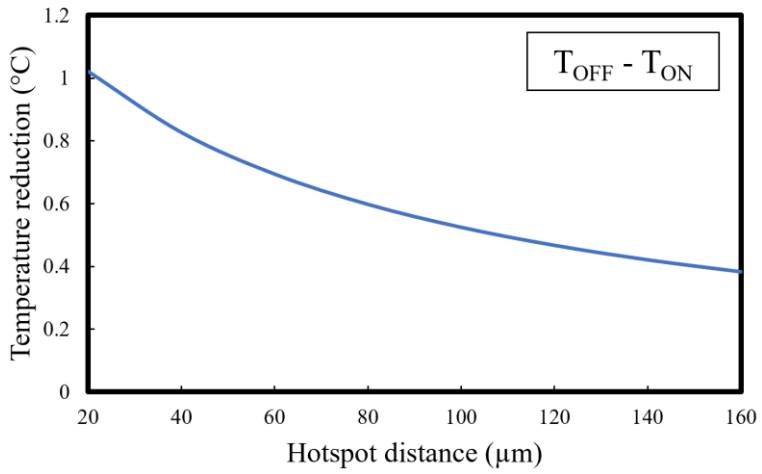

Figure 3-5. Relationship between temperature reduction and hotspot distance.

3.1.4 Effect of TEC length

Another important geometrical parameter is the TEC length (b), which has an impact on the device footprint and the fabrication cost. Similar to the hotspot distance study, we perform a double parameter-sweep in terms of the TEC length (b) and the TEC current. Table 3-3 demonstrates the specific parameters utilized in this simulation:

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	20 µm
b	**TEC length**	**Varies 20-200 µm**
c	Length (x) of cooler	1µm
d	Length (x) of ground	1µm
e	Thickness (z) of chip	10µm
f	Distance between hotspot and top surface	1µm
L_{HS}	Length (x) of hotspot	1µm
T_{HS}	Thickness (z) of hotspot	1µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	1000 W/m^{2}K
Q	Hotspot heat flux	100 W/cm^2
T_∞	Ambient temperature	50 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	2.6 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	35 W/mK
k_{Si}	Thermal conductivity of bulk silicon	130 W/mK
k_{Cu}	Thermal conductivity of copper	385 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-3. Specific parameters used for TEC length study.

As shown in Figure 3-6, the temperature reduction exhibits a positive correlation with the TEC length. A sufficiently long TEC length helps to mitigate Fourier heat conduction, which adversely affects the overall cooling performance. This effect is particularly pronounced when an insulated or convective boundary is applied at the hot end. However, as the heat sink becomes more efficient (i.e., with a larger heat convection coefficient), this length effect will become less significant. Hence, a high-performance heat sink is imperative to maintain space efficiency in the lateral holey-silicon TEC.

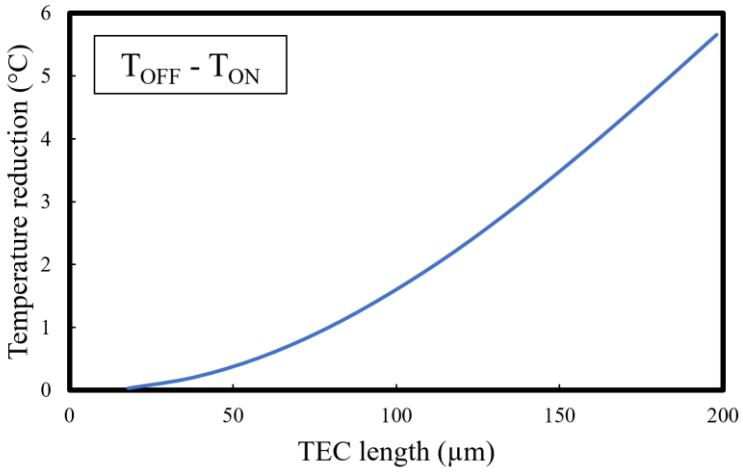

Figure 3-6. Relationship between temperature reduction and TEC length.

3.1.5 Effect of ambient temperature

The ambient temperature plays an essential role in TEC cooling performance, as the thermoelectric figure of merit, $ZT=S^2\sigma T/k$, is proportional to the absolute temperature. With a double parameter-sweep regarding the ambient temperature (T_{amb}) and the TEC current, we evaluate the TEC performance in various ambient temperatures. Table 3-4 demonstrates the specific parameters utilized in this simulation:

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	20 µm
b	TEC length	80 µm
c	Length (x) of cooler	1µm
d	Length (x) of ground	1µm
e	Thickness (z) of chip	10µm
f	Distance between hotspot and top surface	1µm
L_{HS}	Length (x) of hotspot	1µm
T_{HS}	Thickness (z) of hotspot	1µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	1000 W/m^{2}K
Q	Hotspot heat flux	100 W/cm^2
T_∞	**Ambient temperature**	**Varies 50-200 °C**
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	2.6 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	35 W/mK
k_{Si}	Thermal conductivity of bulk silicon	130 W/mK
k_{Cu}	Thermal conductivity of copper	385 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-4. Specific parameters used for ambient temperature study.

The data presented in Figure 3-7 demonstrates that an elevated ambient temperature can result in a greater reduction in hotspot temperature. Furthermore, the linear nature of the temperature reduction curve with respect to the ambient temperature aligns perfectly with the definition of thermoelectric efficiency, thereby confirming the accuracy of our model.

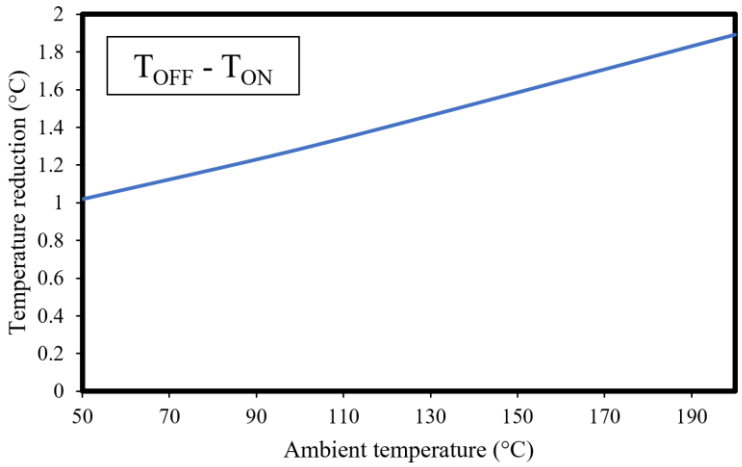

Figure 3-7. Relationship between temperature reduction and ambient temperature.

3.1.6 Effect of hotspot heat flux

The hotspot power is a parameter that can be readily quantified by the transistor power dissipation, therefore knowing its relationship with the TEC cooling performance can provide useful design and operational instructions. The following table shows the specific parameters used the simulation, where a double parameter-sweep regarding the input power density (Q) and the TEC current is conducted:

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	20 µm
b	TEC length	80 µm
c	Length (x) of cooler	1µm
d	Length (x) of ground	1µm
e	Thickness (z) of chip	10µm
f	Distance between hotspot and top surface	1µm
L_{HS}	Length (x) of hotspot	1µm
T_{HS}	Thickness (z) of hotspot	1µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	1000 W/m^{2}K
Q	**Hotspot heat flux**	**Varies 50-500 W/cm^2**
T_∞	Ambient temperature	50 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	2.6 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	35 W/mK
k_{Si}	Thermal conductivity of bulk silicon	130 W/mK
k_{Cu}	Thermal conductivity of copper	385 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-5. Specific parameters used for heat flux study.

As shown in Figure 3-8, the trend of temperature reduction regarding hotspot heat flux exhibits a positive linear relationship, similar to the trend observed in the ambient temperature. At this point, a higher hotspot heat flux leads to a larger reduction in temperature between TEC OFF and TEC ON, which is advantageous for high-power devices.

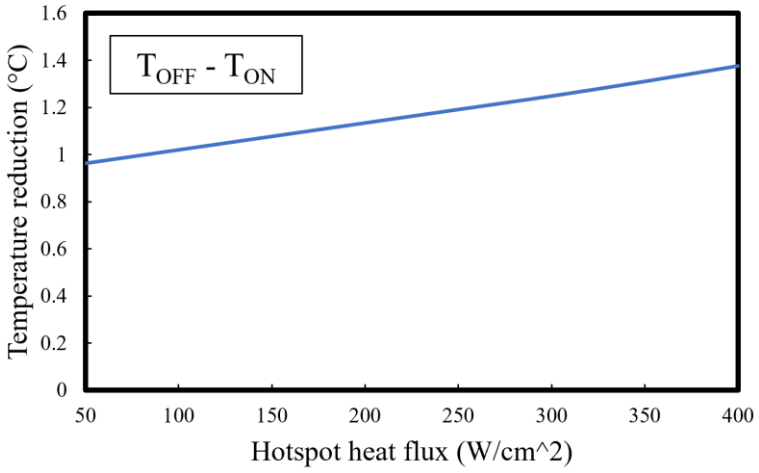

Figure 3-8. Relationship between temperature reduction and hotspot heat flux.

3.1.7 Summary of 2D parametric study

Table 3-6 summarizes the TEC cooling performance trend regarding 14 important design parameters, including geometrical variables (blue), material properties (green) and boundary conditions (red). Above all, the parametric study provides important guidance for future design and optimization of the transistor-TEC system.

Parameter	Expression	Parameter Direction	TEC Cooling Performance
Hotspot distance	a	↑	↓
TEC length	b	↑	↑
Cooler length	c	↑	↑
Heater length	d	↑	↑
Chip thickness	e	↑	↓
distance between top surface and hot spot	f	↑	-
Total width (with constant proportion)	L=a+b+c+d (a:b:c:d=const)	↑	↑
Holey Si Portion (with constant total width)	b/(a+b+c+d)	↑	↑
Thermal Conductivity of Si	k_{Si}	↑	-
Thermal Conductivity of Holey Si (In-plane)	k_x	↑	↓
Thermal Conductivity of Holey Si (Cross-plane)	k_z	↑	-
Convection Coefficient	h_{eff}	↑	↑
Ambient temperature	T_∞	↑	↑
Hotspot heat flux	Q	↑	↑

Table 3-6. Summary of parametric study.

3.2 3D Simulation

3.2.1 Reference simulation

For more practical and reliable analysis, we perform the 3D simulation considering a square transistor-TEC system. Figure 3-9 shows the schematic of the 3D model. Only a quadrant of the geometry is established in the model due to the symmetric configuration.

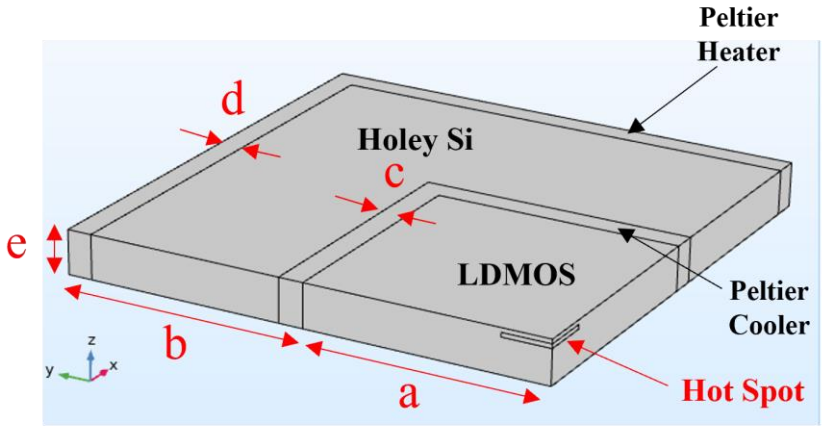

Figure 3-9. Schematic of the 3D transistor-TEC model.

The reference model is pre-optimized based on the parametric study from 2D simulation. Table 3-7 shows the specific parameters used in the simulation. Specifically, the hotspot heat flux is set at high value (i.e., 50700 W/cm^2) to mimic the extreme condition where the potential thermal failure (i.e., thermal runaway) could happen. The effective heat convection coefficient is corrected as 5000 W/m^{2}K based on the reported value in the literature [32], [51]. The thermal conductivity of silicon is modified to 65 W/mK to precisely show the size effect in chip scale. The copper electrodes are replaced by aluminum with thermal conductivity of 237 W/mK. Moreover, the in-plane and cross-plane thermal conductivity of holey silicon are changed to 1 W/mK and 13 W/mK with reduced neck size of 10 nm.

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	50 µm
b	TEC length	50 µm
c	Length (x) of cooler	2.5 µm
d	Length (x) of ground	2.5 µm
e	Thickness (z) of chip	10 µm
f	Distance between hotspot and top surface	1 µm
L_{HS}	Length (x) of hotspot	10 µm
T_{HS}	Thickness (z) of hotspot	1 µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^2K
Q	Hotspot heat flux	50700 W/cm^2
T_∞	Ambient temperature	20 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-7. Specific parameters used for 3D reference simulation.

Figure 3-10 demonstrates the relationship between the hotspot temperature and the TEC current. The optimal current exists at 20 mA, providing a maximum hotspot temperature reduction of 34. 08 °C.

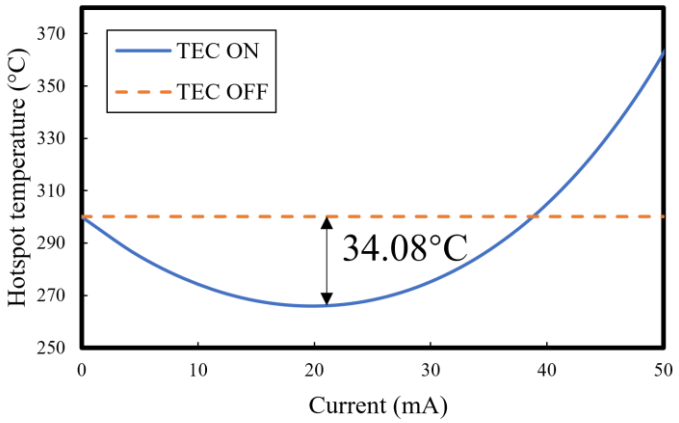

Figure 3-10. Relationship between hotspot temperature and TEC current.

The temperature maps for both TEC OFF and TEC ON are shown in Figure 3-11. In the case of TEC OFF, passive cooling which follows Fourier heat conduction serves as the only cooling mechanism in the transistor-TEC system. As a result, the hotspot temperature is 300 °C

due to excessive power dissipation. On the other hand, when the TEC is ON, active cooling in the TEC aggressively cools down the hotspot temperature to 266 °C.

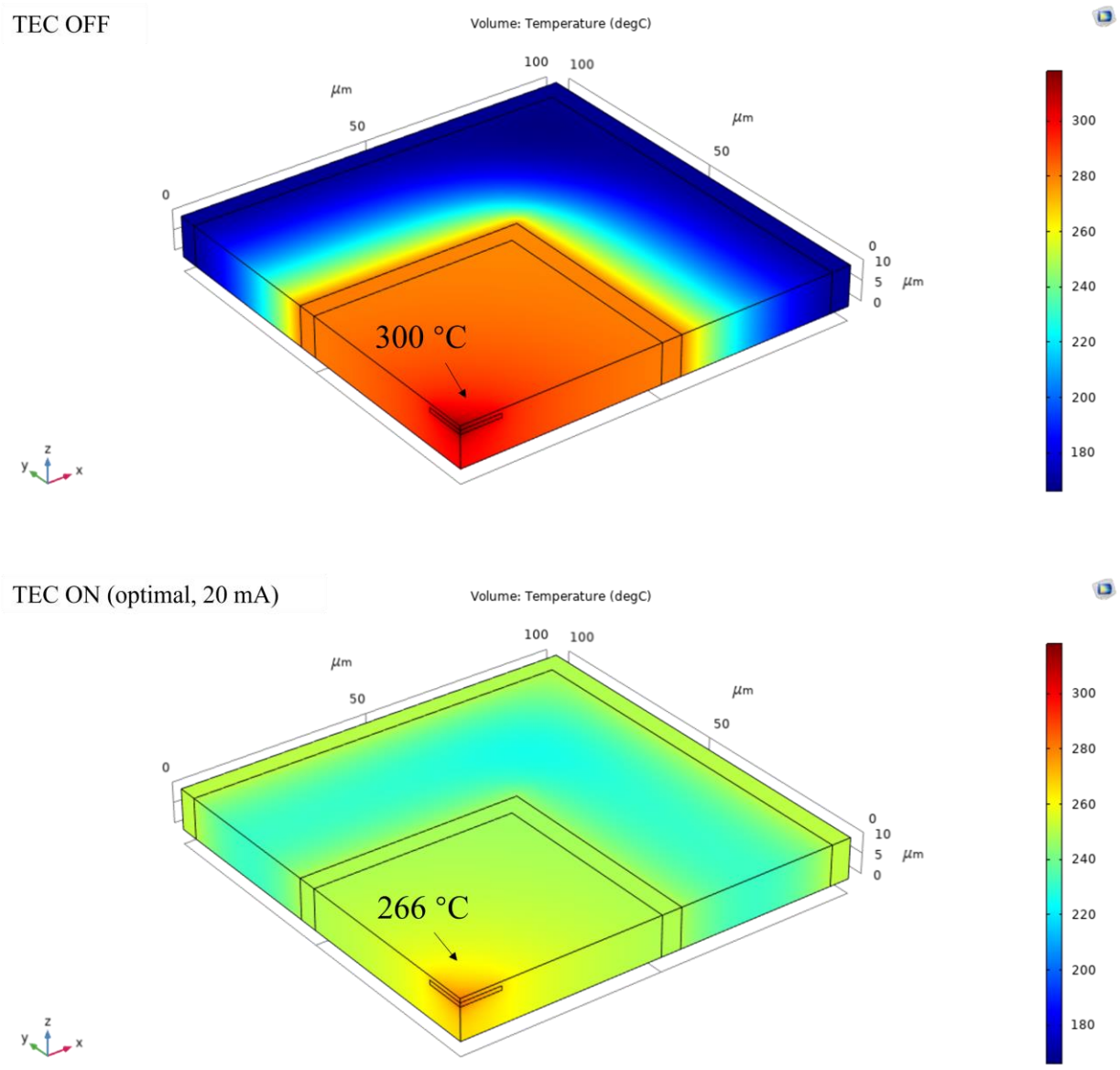

Figure 3-11. Comparison of temperature distribution between TEC OFF and TEC ON with optimal current in 3D reference simulation.

While the 2D reference model provides a reduction of only 1.02 °C in hotspot temperature, the 3D reference model gives up to 34. 08 °C hotspot temperature reduction. Three major factors contribute to the significant improvements of cooling performance: firstly, the enhanced heat convection coefficient ensures more efficient heat removal in the heater electrode. Secondly, the reduced anisotropic thermal conductivity of holey silicon further enhances the ZT value. Finally, high hotspot power facilitates the TEC to provide greater temperature reduction.

3.2.2 Multi-parameter optimization with fixed chip size

The lateral chip space is a critical design constraint in practical applications, necessitating the optimization of TEC area allocation to achieve sufficient cooling within a given chip size. Table 3-8 presents the specific parameters used in this optimization process, with the independent variable highlighted in red and the dependent variables indicated in blue. The specific chip size is fixed at 100 × 100 μm^2.

Geometrical parameters

Parameters	Name	Value
a	**Hotspot distance**	**Varies 20-80 μm**
b	TEC length	100-a
c	Length (x) of cooler	0.05 × b
d	Length (x) of ground	0.05 × b
e	Thickness (z) of chip	10 μm
f	Distance between hotspot and top surface	1 μm
L_{HS}	Length (x) of hotspot	10 μm
T_{HS}	Thickness (z) of hotspot	1 μm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^2K
Q	Hotspot heat flux	50700 W/cm^2
T_∞	Ambient temperature	20 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 μV/K

Table 3-8. Specific parameters used in the optimization with fixed chip size.

As depicted in Figure 3-12, the increased TEC area can be achieved by sacrificing the transistor area, resulting in a significantly enhanced cooling performance. Specifically, when combined with a 20 μm hotspot distance and an 80 μm TEC length, it enables a temperature reduction of 121.6 °C. This remarkable cooling performance is attributed to both the close proximity between the hotspot and cooler as well as the improved performance due to the extended TEC length. A more detailed explanation can be found in the 2D parametric study presented in section 3.1.

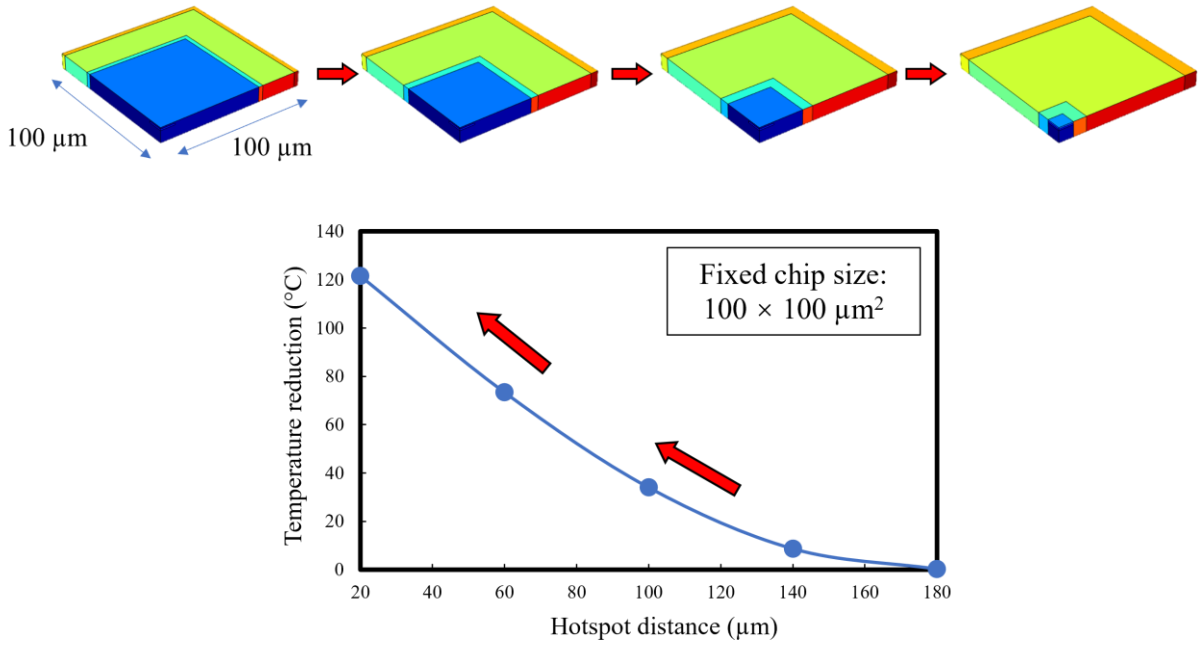

Figure 3-12. Optimization results with fixed chip size. (Top) schematic of device optimization under fixed chip size. (Bottom) relationship between temperature reduction and hotspot distance with fixed chip size.

3.2.3 Multi-parameter optimization with fixed a/b ratio

Another critical design constraint lies in the transistor area, which plays an essential role in the electrical characteristics of the transistor device. The optimization of transistor area can be done by changing the TEC length with a given a/b ratio. Table 3-9 specifies the parameters used in this optimization where the transistor size is varied from $(1\ \mu m)^2$, $(10\ \mu m)^2$, $(100\ \mu m)^2$ to $(200\ \mu m)^2$ and the a/b ratio is changed from 1:2, 1:1, 2:1 to 5:1.

Geometrical parameters

Parameters	Name	Value
a	**Hotspot distance**	**Varies 0.5-100 µm**
b	**TEC length**	**Varies 0.1-200 µm**
c	Length (x) of cooler	0.05 × b
d	Length (x) of ground	0.05 × b
e	Thickness (z) of chip	10 µm
f	Distance between hotspot and top surface	1 µm
L_{HS}	Length (x) of hotspot	10 µm
T_{HS}	Thickness (z) of hotspot	1 µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^2K
Q	Hotspot heat flux	50700 W/cm^2
T_∞	Ambient temperature	20 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 3-9. Specific parameters used in the optimization with fixed a/b ratio.

In the previous study, we concluded that a smaller hotspot distance (i.e., transistor size) helps to improve the cooling performance. However, as shown in Figure 3-13, with a fixed a/b ratio, the cooling performance is increased by increasing the hotspot distance. The contradiction in these conclusions lies in the fact that the TEC length is a more impactful independent variable compared to the hotspot distance in terms of TEC cooling performance. Without the constraint in chip size, one should always consider increasing the TEC length to maximize the hotspot temperature reduction.

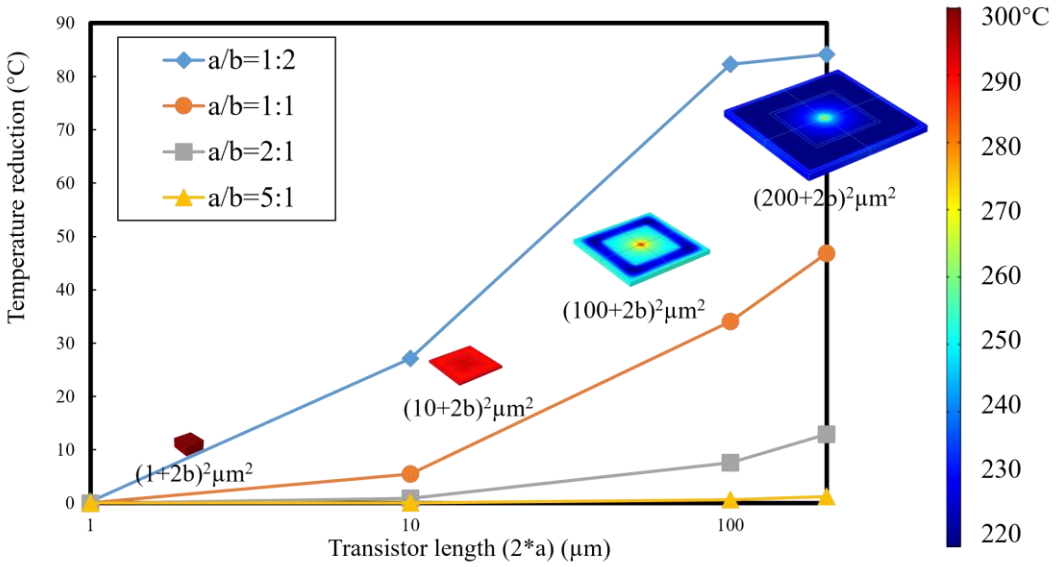

Figure 3-13. Optimization results with fixed a/b ratio. Four values of a/b ratio, 1:2, 1:1, 2:1 and 5:1, are studied with transistor size of (1 μm)², (10 μm)², (100 μm)² and (200 μm)²

3.3 1D Analytical Model and TEC Advantage

In section 3.1 and 3.2, the TEC cooling performance is represented by the hotspot temperature reduction (i.e., the difference between the hotspot temperature in OFF state and the hotspot temperature in ON state). Although significant enhancements in TEC cooling performance have been realized through a series of optimization, the nature of low in-plane thermal conductivity of holey silicon makes it potentially less competitive against the high thermal conductivity bulk silicon-based system which implements only passive cooling. To compare the TEC cooling with the passive cooling mechanism, we investigate the 1D analytical model and study the TEC advantage against the bulk silicon-based system.

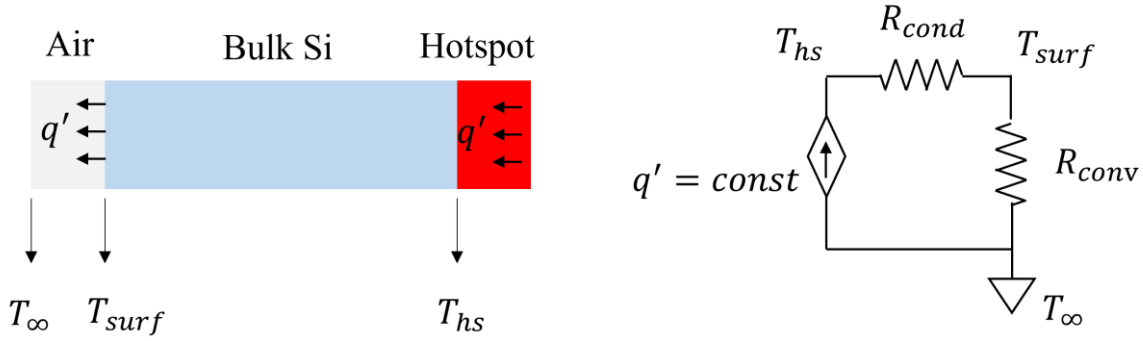

Figure 3-14. 1D heat transfer model with single passive cooling. (Left) The schematic. (Right) The equivalent thermal circuit.

Figure 3-14 depicts the 1D heat transfer model with single passive cooling. A constant heat flux is assigned in the hot end (i.e., the hotspot region) while a convective boundary is designated at the cold end (i.e., the heat sink). The hotspot heat flux and temperature can be expressed by:

$$q' = \frac{k(T_{hs} - T_{surf})}{L} = h(T_{surf} - T_\infty) \tag{17}$$

$$T_{hs} = \frac{q'L}{k} + \frac{q'}{h} + T_\infty \tag{18}$$

where q' is the hotspot heat flux, k is the thermal conductivity, h is the convection coefficient, T_{hs} is the hotspot temperature, T_{surf} is the surface temperature, T_∞ is the ambient temperature and L is the length of the conductive material (e.g., bulk silicon).

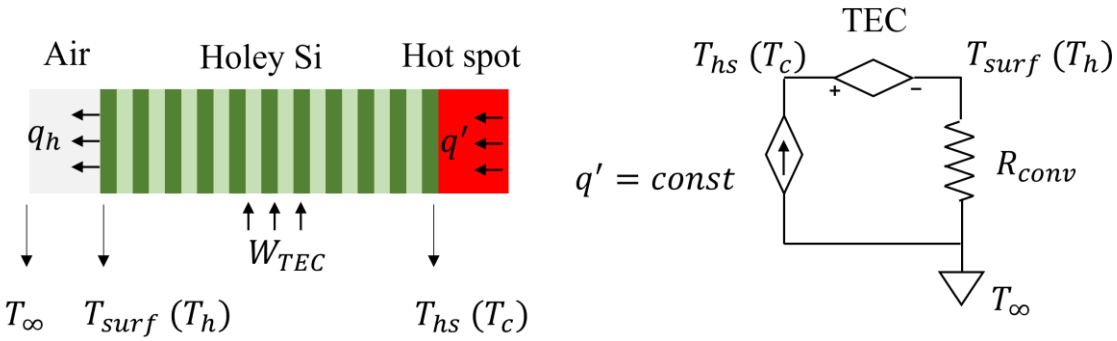

Figure 3-15. 1D heat transfer model with TEC cooling. (Left) The schematic. (Right) The equivalent thermal circuit.

The second equation above indicates a linearity between the hotspot temperature and the ambient temperature with a slop of 1. Besides, a reduced hotspot temperature can be obtained by increasing the thermal conductivity of the conductive material.

Figure 3-15 illustrates the 1D heat transfer model with TEC cooling. The boundary conditions are considered the same as the above passive cooling model. In this case, the equations can be written as:

$$q' = q_c = ST_cI - \frac{\Delta T}{R_{th}} - \frac{I^2R}{2} \tag{19}$$

$$W_{TEC} = q_c - q_h = S\Delta TI + I^2R \tag{20}$$

$$q_h = ST_hI - \frac{\Delta T}{R_{th}} + \frac{I^2R}{2} = h(T_h - T_\infty) \tag{21}$$

$$T_c = \frac{(h - SI)\left(q' + \frac{I^2R}{2}\right)R_{th} + I^2R + q'}{SIR_{th}(h - SI) + h} + \frac{1}{\frac{SIR_{th}(h - SI)}{h} + 1}T_\infty \tag{22}$$

where q' is the hotspot heat flux, q_c is the TEC cooling flux, S is the Seebeck coefficient, R_{th} is the thermal resistance, T_c (T_{hs}) is the cold side (hotspot) temperature, T_h (T_{surf}) is the hot side

(surface) temperature, ΔT is the temperature difference ($\Delta T = T_h - T_c$), L is the length of the TEC, I is the TEC current, R is the electrical resistance, W_{TEC} is the work of TEC, q_h is the TEC heating flux, h is the convection coefficient and T_∞ is the ambient temperature.

Similar to the passive cooling scenario, the hotspot temperature is linear to the ambient temperature. However, the slope here is $\frac{1}{\frac{SIR_{th}(h-SI)}{h}+1}$, which is a comprehensive variable dependent on both the material properties and the boundary conditions. If $h < SI$, the slope is less than 1, otherwise the slop is equal or larger than 1. At this point, Figure 3-16 shows the TEC advantage zone where the hotspot temperature in the transistor-TEC system is lower than that in the bulk silicon-based system:

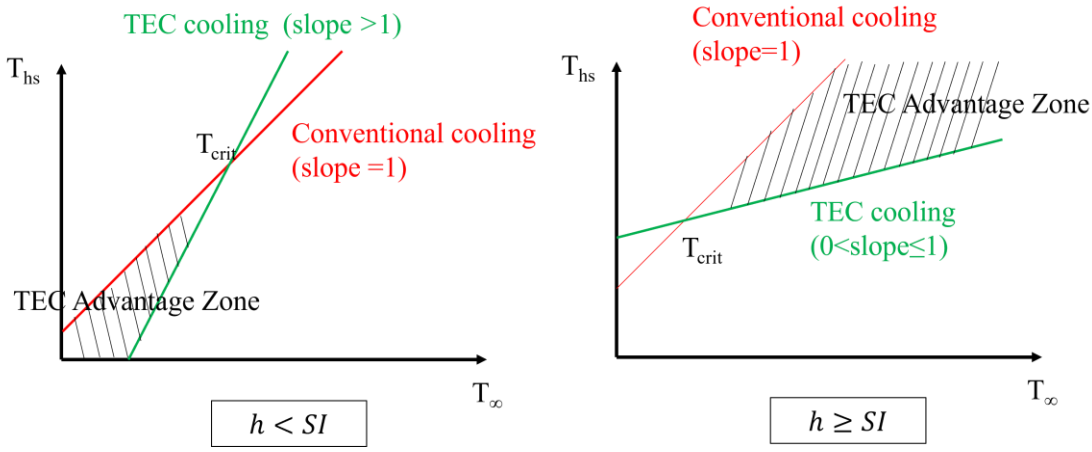

Figure 3-16. TEC Advantage zone. (Left) $h < SI$. (Right) $h \geq SI$.

When the heat convection coefficient is small, the transistor-TEC system can provide a lower hotspot temperature compared to the bulk silicon-based system when the ambient temperature is also small. Once the ambient temperature exceeds a certain value, the advantage of TEC no longer exists. On the other hand, if the heat convection coefficient is sufficiently large, the transistor-TEC system will achieve a lower hotspot temperature after surpassing a critical

ambient temperature. Nevertheless, determining this critical temperature (i.e., intersection point) depends on various parameters related to both the material properties and the boundary conditions (i.e., the first term of T_{hs}). Generally, reduced k, reduced R and increased S always contribute to expanding the TEC advantage zone; however, the changes in k_{TEC}, R_{th}, I and h provide a collective effect.

The limitation of TEC advantage zone makes TEC not always preferrable for steady-state cooling. A careful evaluation of the heating conditions and geometrical parameters should be done before implementing the TEC as a cooling solution.

Chapter 4

Transient TEC Cooling

In microelectronic systems, the heating conditions may undergo dynamic evolution due to device aging, changing electrical characteristics, different external environments and various user operations. While the steady-state TEC simulation provides primary guidelines for optimization, the transient TEC simulation offers more realistic and accurate analysis. Additionally, transient TEC cooling has its unique behavior which can benefit the microelectronic cooling. Specifically, transient TEC cooling exhibits spatial and temporal mismatches between the Peltier effect and the concomitant Joule effect. While the cooling occurs at the cooler junction simultaneously with TEC activation, the volumetric heating takes place throughout the TEC with a delayed response due to heat diffusion.

Different from the steady-state TEC cooling, the governing equation in transient TEC cooling incorporates the time variable, which can be written as:

$$\nabla \cdot \boldsymbol{q} = Q_e + \rho C_p \frac{dT}{dt} \tag{21}$$

The equations (18), (19) and (20) utilized to describe the thermoelectric effect remain unchanged, with the exception that all variables are now time dependent.

4.1 Hotspot with Constant Heat Flux

The transient TEC simulation is first performed with a stationary hotspot (i.e., with constant heat flux). The 2D cross-sectional sketch and 3D geometry used in the transient TEC model remains unchanged, which are demonstrated in Figure 4-1. Specific parameters used in the simulation, including geometrical parameters, material properties and boundary conditions, are demonstrated in detail in Table 4-1.

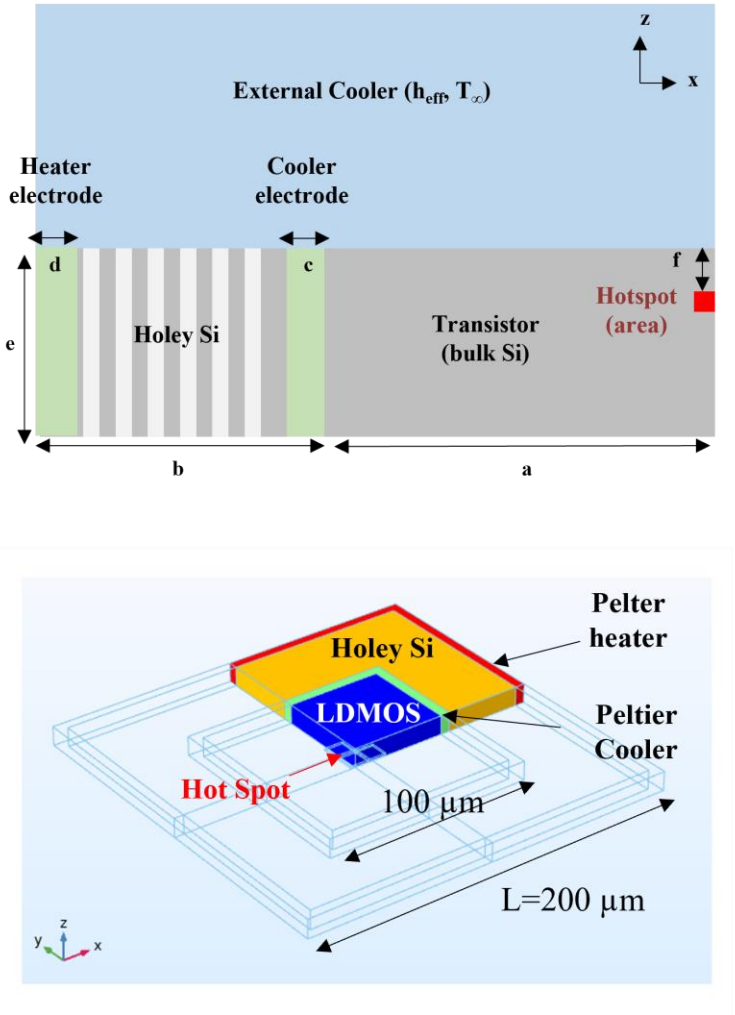

Figure 4-1. Schematics of transient TEC model. (Top) A 2D cross-sectional sketch. (Bottom) The 3D geometry implemented in the simulation.

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	50 µm
b	TEC length	50 µm
c	Length (x) of cooler	5 µm
d	Length (x) of ground	5 µm
e	Thickness (z) of chip	10 µm
f	Distance between hotspot and top surface	1 µm
L_{HS}	Length (x) of hotspot	10 µm
T_{HS}	Thickness (z) of hotspot	1 µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^{2}K
Q	Hotspot heat flux	50700 W/cm^2
T_∞	Ambient temperature	20 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 4-1. Specific parameters used in the transient simulation with constant heat flux.

Unlike the steady-state TEC cooling which utilizes a constant TEC current, in transient TEC cooling, a time-dependent TEC current is applied. Two patterns of transient current is shown in Figure 4-2 with the corresponding cooling results:

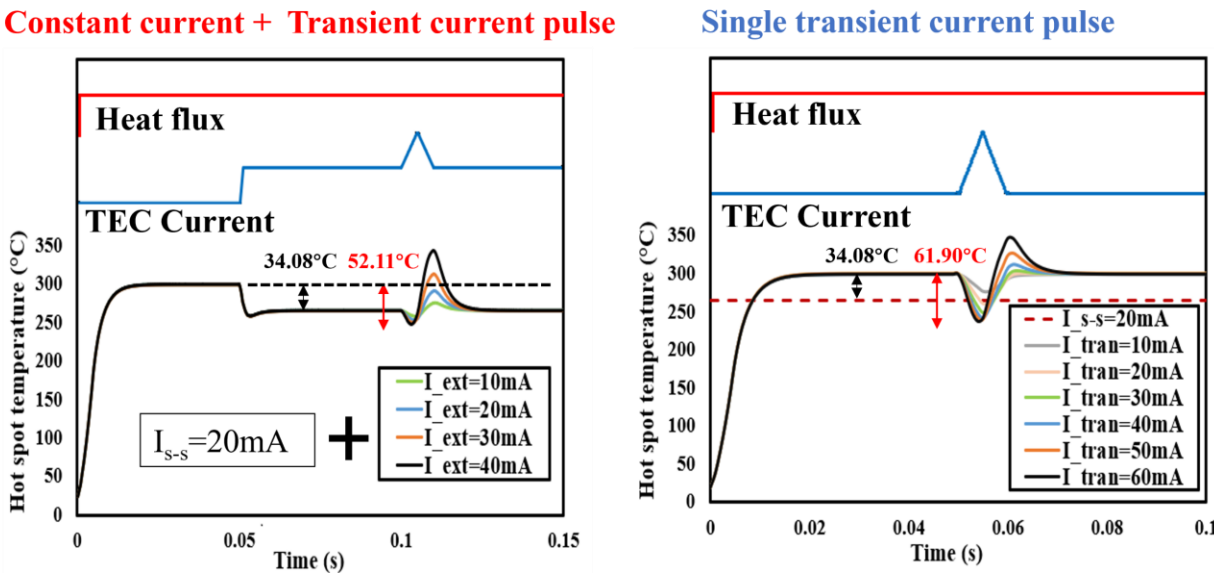

Figure 4-2. Two patterns of transient current and their corresponding cooling results. (Left) constant current + transient current pulse. (Right) single transient current pulse.

4.1.1 Constant current + transient current pulse

In the first case of Figure 4-2 (left), a constant heat flux with a power density of 50700W/cm^2 is applied to heat the hotspot up to 300 °C and achieve the first steady state. Subsequently, a constant TEC current of 20 mA is employed to attain the second steady state. This specific current magnitude has been optimized in the previous study (i.e., section 3.2.1) to ensure maximum reduction in steady-state hotspot temperature (i.e., 34.1 °C). Following this, an additional transient current pulse is introduced alongside the constant current. The amplitude of this extra transient current pulse ranges from 10 mA to 40 mA and the duration is fixed at 10 ms.

With the TEC current pulse, the temperature further drops down to 247.9 °C with a temperature reduction of 52.1 °C. Later, the hotspot undergoes a temperature overshot, which heats up the hotspot to over 300 °C. This additional cooling in transient state is referred to as the supercooling effect [20], [29], [30], [67].

4.1.2 Single transient current pulse

Another case shown in Figure 4-2 (right) is to implement a single transient current pulse. While the constant heat flux remains the same, the constant TEC current is no longer applied in this scenario. Instead, a single transient current pulse is employed with an amplitude ranging from 10 mA to 60 mA. The pulse duration is fixed at 10 ms.

Similar to the former case, when the transient current pulse is applied, the transistor-TEC system exhibits a sudden temperature drop, followed by a delayed temperature overshot. More importantly, the maximum reduction in hotspot temperature is now increased to 61.9 °C. This superior cooling performance may happen because of the following reasons: first, the patterns of

the current pulse are different. The ramp-up slope for the second triangular pulse is much larger than the first one. Second, the Joule heating effects are different. The former case has reached the steady-state Joule heating beforehand, while the latter case experienced transient Joule heating during the entire TEC activation.

The single transient current pulse has greater cooling performance compared to the combination of a constant current and a transient current pulse, a fact that makes it more promising in the actual applications. To provide further optimization in the case of single transient current pulse, Figure 4-3 demonstrates the peak hotspot temperature during early supercooling and sequential temperature overshoot. As the transient TEC current increases, the early supercooling results in reduced hotspot temperatures. This hotspot temperature can be even lower than the minimum temperature in steady-state cooling with optimal constant current. However, when the transient current magnitude is sufficiently large, the hotspot temperature will reach a plateau. On the other hand, the sequential temperature overshoot exhibits a exponent trend, where the hotspot

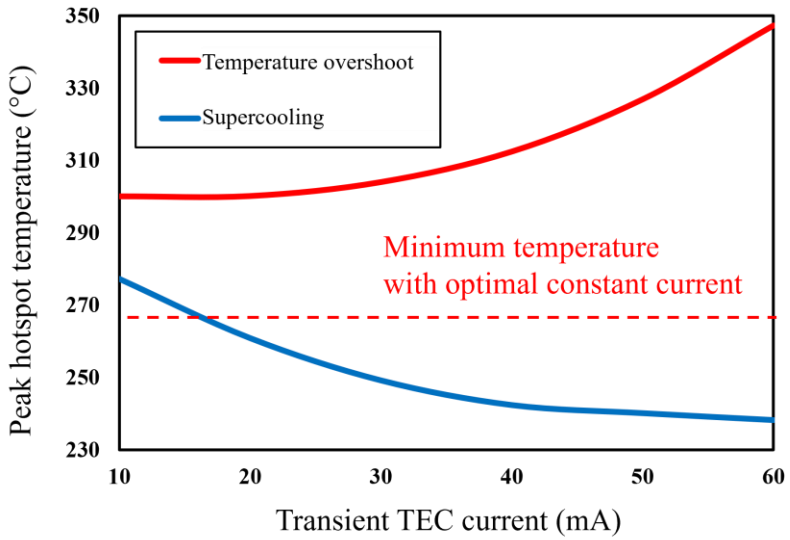

Figure 4-3. Peak hotspot temperature in early supercooling (blue) and sequential temperature overshoot (red).

temperature will keep increasing as the transient TEC increases. Therefore, the implementation of such single transient current pulse should consider its threshold in temperature reduction as well as its side effect of temperature overshoot.

4.2 Hotspot with Transient Heat Flux

4.2.1 Reference simulation

When dealing with constant heat flux, the transient TEC cooling experiences early supercooling as well as delayed temperature overshot, as discussed in section 4.1. This temperature overshoot exceeds the steady-state temperature, creating additional challenges in thermal management. To alleviate this adverse effect, here we introduce the idea of using transient TEC cooling to deal with transient heat flux, as is demonstrated in Figure 4-4: without TEC activation, a transient heat pulse with a power density of 50700 W/cm^2 and a duration of 5 ms raises the hotspot temperature from 20 ℃ to 226.4 ℃ (black dashed curve); however, a simultaneous TEC current pulse with a magnitude of 60 mA and a duration of 10 ms can aggressively reduce the hotspot temperature to 158.5 ℃, showing a temperature reduction of 67.9 ℃ (orange solid curve) without significant temperature overshoot.

Different from the constant heat flux scenario, the temperature overshoot is now offset by the sudden removal of the transient heat flux. A careful optimization process should be performed so that the moment when the transient hotspot deactivates can align well with the moment when the temperature overshoot begins. In the upcoming simulation, such fitting is done by optimizing the conditions of the transient TEC current pulse (e.g., amplitude, location, duration and pulse shape), which will be discussed in detail in the following sections. Specific parameters used in the transient optimization can be found in Table 4-1.

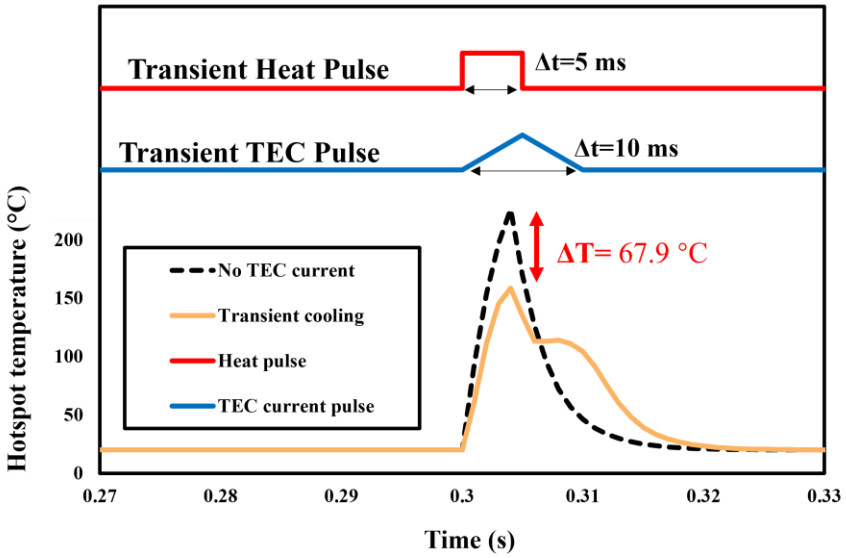

Figure 4-4. Concept of using transient TEC cooling to deal with transient heat flux.

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	50 μm
b	TEC length	50 μm
c	Length (x) of cooler	5 μm
d	Length (x) of ground	5 μm
e	Thickness (z) of chip	10 μm
f	Distance between hotspot and top surface	1 μm
L_{HS}	Length (x) of hotspot	10 μm
T_{HS}	Thickness (z) of hotspot	1 μm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^{2}K
Q	Hotspot heat flux	50700 W/cm^2
T_∞	Ambient temperature	100 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 μV/K

Table 4-2. Specific parameters used in the transient simulation with time-dependent heat flux.

4.2.2 Amplitude-dependent cooling performance

The TEC current amplitude plays a crucial role in transient TEC cooling. In the following simulation, this amplitude is regarded as the only variable for cooling performance evaluation. To

simulate extreme heating conditions that can potentially lead to thermal failure (i.e., thermal runaway), a hotspot heat pulse with a power density of 50700 W/cm^2 is employed with a high background temperature of 100 °C. In response to this high input heat flux, a TEC current pulse is then implemented as the active cooling method in the transistor-TEC system. The duration of the heat pulse is 5 ms, while that of the TEC current pulse is 10 ms.

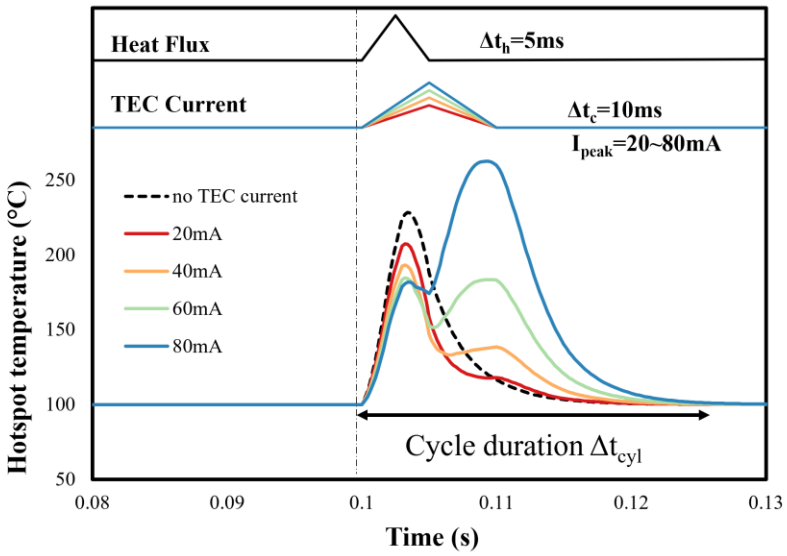

Figure 4-5. Evolution of hotspot temperature under various TEC current amplitudes.

Figure 4-5 shows the cooling performance of the transistor-TEC device under various TEC current amplitudes. Without TEC activation, the transient heat pulse results in a peak hotspot temperature of 228.3 °C in 3.2 ms. Following this, an increase in TEC current amplitude helps to reduce this peak hotspot temperature. However, this increasing TEC current amplitude also leads to more significant effect of delayed temperature overshoot, leading to a second peak hotspot temperature after the transient heat flux ends. Apparently, the higher value of these two peaks determines the maximum hotspot temperature, therefore a moderate amplitude is required to balance these two sequential heating effects. Table 4-3 records the hotspot temperature in the first

and second peak. In conclusion, a TEC current amplitude of 60 mA provides the best cooling performance.

TEC current amplitude	***First peak temperature***	***Second peak temperature***	***Temperature reduction***
No TEC current	228.3℃	-	-
20mA	207.3℃	118.0℃	21.0℃
40mA	193.3℃	138.5℃	35.0℃
60mA	184.6℃	183.5℃	43.8℃
80mA	181.9℃	262.6℃	No cooling

Table 4-3. Peak hotspot temperature under different TEC current amplitudes.

4.2.3 Location-dependent cooling performance

The initiation moment of the transient cooling pulse is a crucial variable as it determines the onset time of the early supercooling and the sequential temperature overshoot. In the following simulation, the transient heat pulse with an amplitude of 60 mA and a duration of 10 ms will be initiated in various temporal positions relative to the hotspot heat flux.

The hotspot temperature evolution under a TEC current pulse with different temporal positions is illustrated in Figure 4-6, where the corresponding peak temperature data is shown in Table 4-4. Specifically, initiating the TEC current pulse ahead of the hotspot heat pulse may result in inadequate cooling due to the overlapping effect of the hotspot heat pulse and temperature overshoot. Conversely, initiating the TEC current pulse after the hotspot heat pulse also leads to insignificant temperature reduction due to the temporal mismatch between hotspot heating and

TEC supercooling. Only when the TEC current pulse coincides with the hotspot heat pulse can we achieve a minimum value for peak hotspot temperature.

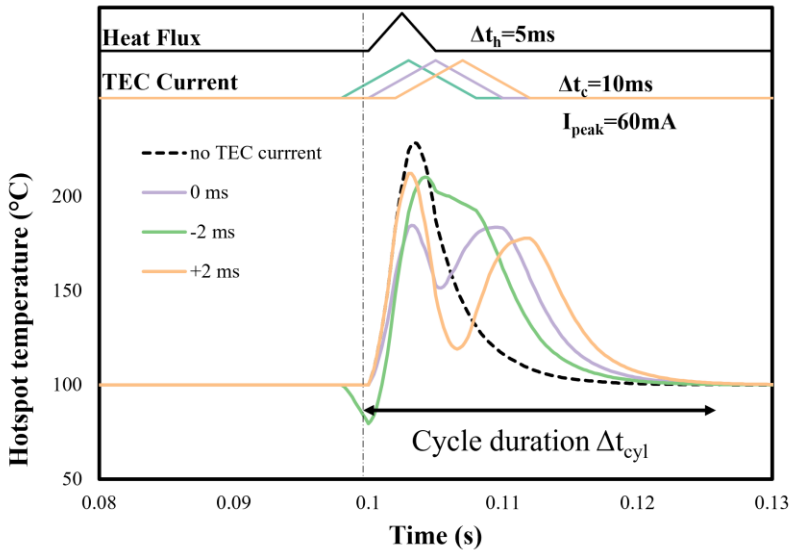

Figure 4-6. Evolution of hotspot temperature under various temporal locations of TEC current pulse.

TEC pulse starting time	***First peak temperature***	***Second peak temperature***	***Temperature reduction***
No TEC current	*228.3°C*	-	-
0ms	*184.6°C*	*183.5°C*	*43.8°C*
-2ms	*210.1°C*	-	*18.2°C*
+2ms	*212.0°C*	*177.8°C*	*16.3°C*

Table 4-4. Peak hotspot temperature under different temporal locations of the TEC current pulse.

4.2.4 Duration-dependent cooling performance

In the previous simulation with constant heat flux, the duration of the TEC current pulse (i.e., 10 ms) is about twice as large as the supercooling period (i.e., ~5ms). Following this rule, a changing TEC current pulse duration may help to control the period of the early supercooling, therefore providing a better match to the hotspot heat pulse. To study the optimal duration of the TEC current pulse, Figure 4-7 demonstrates the hotspot temperature change in time under various durations in TEC current. While the duration of the hotspot heat pulse is fixed at 5ms, the TEC current duration is varied from 3 to 18 ms.

As shown in Figure 4-7, a short duration of the TEC current pulse provides limited supercooling time. This also results in an early temperature overshoot overlapping with the original hotspot heat pulse. On the other hand, a long TEC current pulse duration provides insufficient cooling due to the slowly ramping-up profile. A perfect balance can be found when the duration of the TEC current pulse is about 1.8 times of the hotspot heat pulse. Table 4-5 shows the peak temperature data obtained from the simulation.

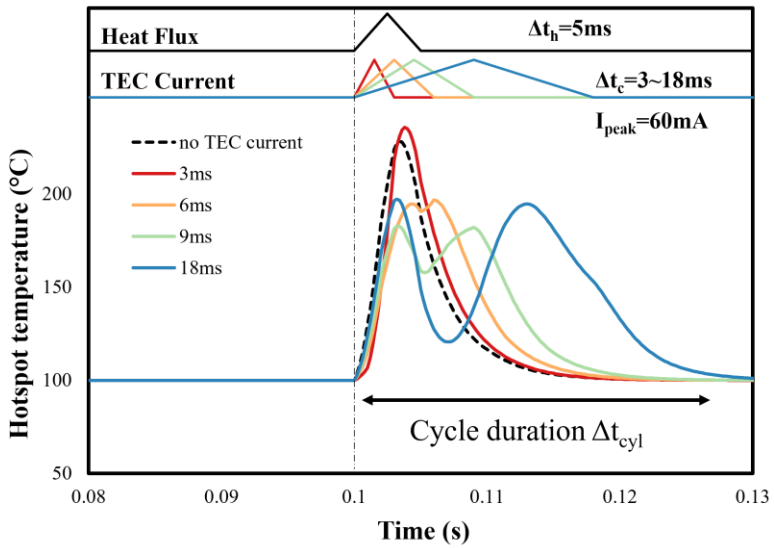

Figure 4-7. Evolution of hotspot temperature under various temporal locations of TEC current.

TEC pulse starting time	*First peak temperature*	*Second peak temperature*	*Temperature reduction*
No TEC current	*228.3°C*	-	-
3ms	*235.8°C*	-	*No cooling*
6ms	*194.7°C*	*196.9°C*	*31.5°C*
9ms	*182.8°C*	*182.1°C*	*45.5°C*

Table 4-5. Peak hotspot temperature under different durations of the TEC current pulse.

4.2.5 Shape-dependent cooling performance

Previous studies have investigated the shape effect in the transient TEC current pulse [20], [28], [68]. However, those papers are limited to the constant heat flux, whereas a dynamic heat pulse is never considered. Here, we perform the shape study of the TEC current pulse under transient heat flux conditions. Four shapes of the TEC current pulse, isosceles triangle, ramp-up triangle, ramp-down triangle and rectangle will be discussed.

As shown in Figure 4-8, two sets of simulations: fixed parameters and optimal parameters are performed. For the simulation with fixed parameters, the hotspot temperature profile shows significant differences in four shapes. For the ramp-down triangular and rectangular TEC current pulse, the rapid increase in TEC current results in overcooling at the early heating stage (i.e., t = 0.1-0.102 s), while excessive amounts of Joule heating is overlapped with the hotspot heat pulse. On the other hand, for the ramp-up triangular TEC current pulse, the slow increase in TEC current results in limited cooling at the heating stage, providing insignificant cooling performance. On the other hand, the TEC current pulse in isosceles triangle retains the advantages among different pulse shapes, therefore leading to minimum peak hotspot temperature among four cases.

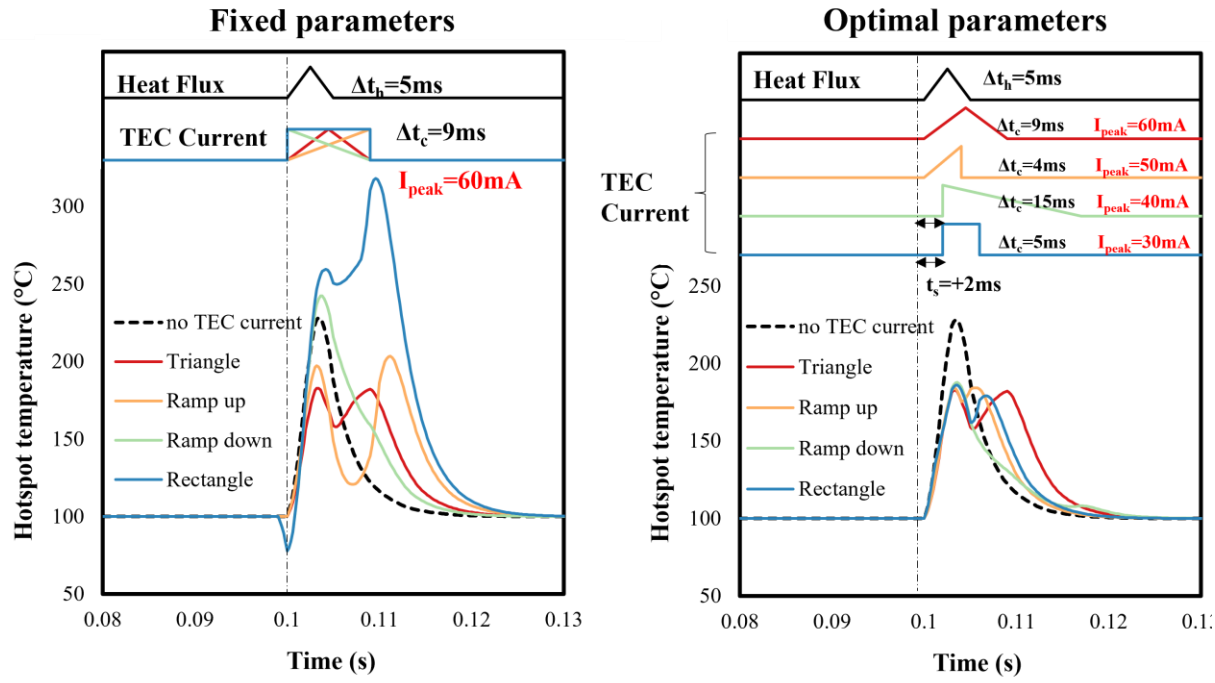

Figure 4-8. Hotspot temperature as a function of time under different TEC current pulse shapes (i.e., isosceles triangle, ramp-up triangle, ramp-down triangle and rectangle) (a) Fixed parameters (i.e., 9ms duration, 60mA amplitude, simultaneous) (b) optimized parameters.

Compared to the simulation with fixed parameters, the simulation with optimal parameters shows similar peak temperature reduction among four pulse shapes. For the ramp-down triangular and rectangular TEC current pulse, the delayed initiation of TEC current pulse compensates for the rapid increase in TEC current. Besides, the optimal duration and amplitude of the ramp-up pulse ensures adequate transient current. The simulation results regarding the reduction in peak hotspot temperature as well as the specific TEC current pulse conditions are summarized in Table 4-6.

TEC current pulse shapes	*Fixed parameters*				*Optimized parameters*			
	Amplitude	*Location*	*Duration*	*Temperature reduction*	*Amplitude*	*Location*	*Duration*	*Temperature reduction*
Isosceles triangle 	60mA	+0ms	9ms	45.5℃	60mA	+0ms	9ms	45.5℃
Ramp-up triangle 				25.0℃	50mA	+0ms	4ms	44.0℃
Ramp-down triangle				-14.0℃ (No cooling)	40mA	-2ms	15ms	40.6℃
Rectangle 				-89.7℃ (No cooling)	30mA	-2ms	5ms	42.2℃

Table 4-6. Summary of the reduction in peak hotspot temperature and the specific TEC current pulse conditions.

Chapter 5

TEC Array Performance

In modern integrated circuit packages, the integration of high-power-density transistor arrays has become a common practice. For instance, in the realm of 5G wireless communication, laterally-diffused metal-oxide-semiconductor (LDMOS) arrays (as shown in Figure 5-1) are widely used in the radio frequency (RF) amplifier to ensure high-quality signal amplification [69], [70]. Similarly, in the field of artificial intelligence (AI) and machine learning, the implementation of multiply accumulate (MAC) arrays in NPUs or TPUs enable a huge volume of parallel mathematical operations [71], [72]. These transistor or transistor-based arrays carry excessive power, leading to localized hotspots that dynamically change in space and time.

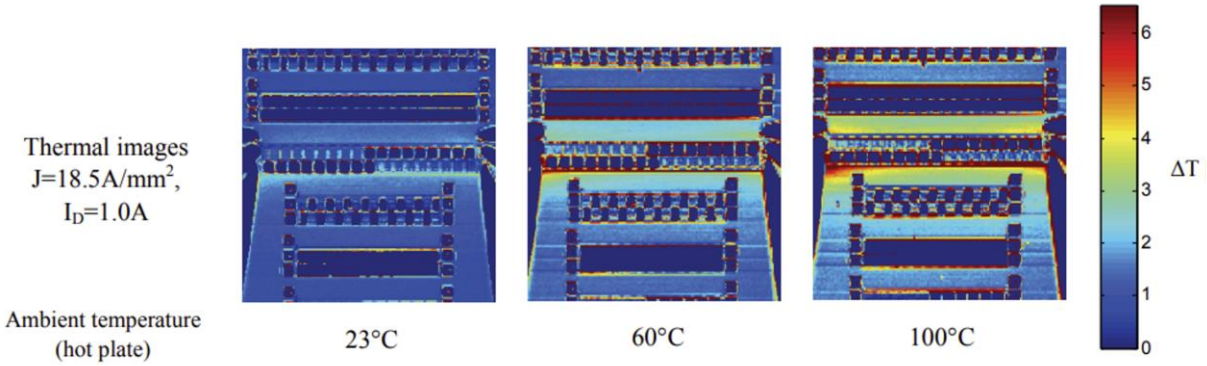

Figure 5-1. Thermal image of LDMOS array under different background temperatures. Figure adapted from [75] .

5.1 Lateral TEC Array Designs

In holey silicon-based lateral TEC, the Peltier heater, Peltier cooler, holey silicon region as well as the cooling object (i.e., transistor) are well aligned in horizontal direction. This special configuration of transistor-TEC system makes it possible for array scaling. Figure 5-2 illustrates three types of lateral TEC array designs. The second column shows the full sketch while the third column shows the actual geometry implemented in the simulation due to symmetry. A 2 × 2 transistor matrix is taken as an example to investigate its coupling effect with the TEC array. Note that the number of rows and columns in a practical transistor-TEC array system are not restricted to the abovementioned values, nor is the number of transistors limited to one for each transistor-TEC unit.

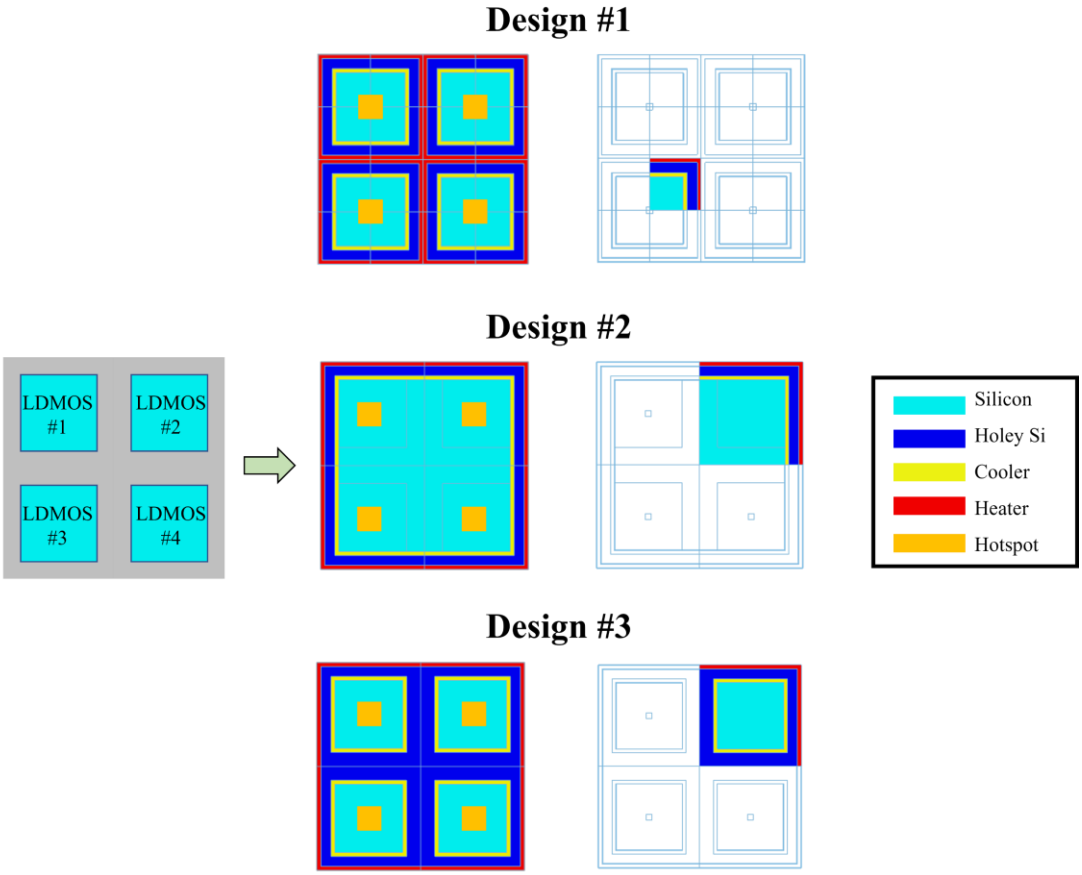

Figure 5-2. Three types of lateral TEC array designs.

Design #1 (Individual TEC): The first design utilizes the direct scaling method, where multiple transistor-TEC systems are connected to each other in lateral direction. In this design, individual TECs can provide selective cooling for separate transistor units, which offers on-demand thermal management to the transistor matrix.

Design #2 (Single TEC): The second design employs one single TEC to cool down the 2 × 2 transistor matrix, which is similar to the single transistor-TEC system. This design is space-efficient as it consumes less space for the TEC. However, it also loses the capability to perform on-demand thermal control.

Design #3 (Individual cooler + single heater): The third design is a combination of the first and second designs, where the individual coolers and single heater are used. By selectively turning on the individual coolers, this array design can provide efficient spatial temperature control under different heating conditions. Moreover, the implementation of a single heater reduces the lateral TEC footprint, making it more suitable for the integration into chip systems with limited space.

5.2 Array Design Optimization

5.2.1 Reference simulation

Figure 5-3 shows the geometry of three different array designs applied in the steady-state reference simulation. All designs share the same parameter set for comparison, as specified in Table 5-1. The simulation results are shown in Figure 5-4, where the hotspot temperature, temperature reduction and temperature maps are shown for both TEC ON and OFF states.

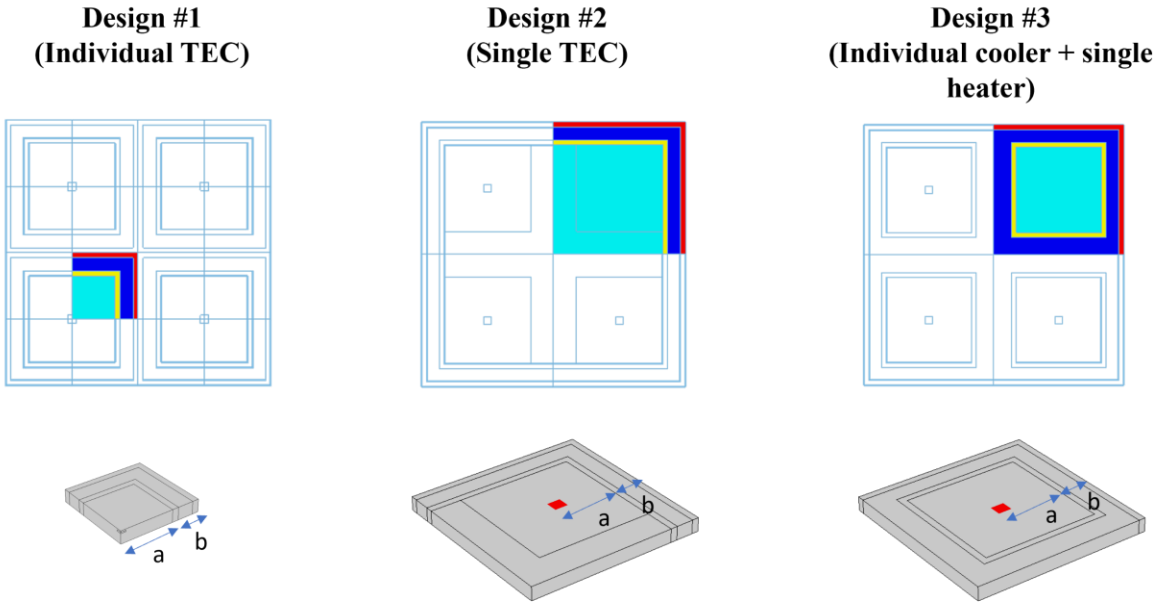

Figure 5-3. Geometry of three different designs in reference simulation.

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	50 µm
b	TEC length	15 µm
c	Length (x) of cooler	1.5 µm
d	Length (x) of ground	1.5 µm
e	Thickness (z) of chip	10 µm
f	Distance between hotspot and top surface	1 µm
L_{HS}	Length (x) of hotspot	4 µm
T_{HS}	Thickness (z) of hotspot	1 µm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^{2}K
Q	Hotspot heat flux	35000 W/cm^2
T_∞	Ambient temperature	20 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 µV/K

Table 5-1. Specific parameters used in the array design simulation.

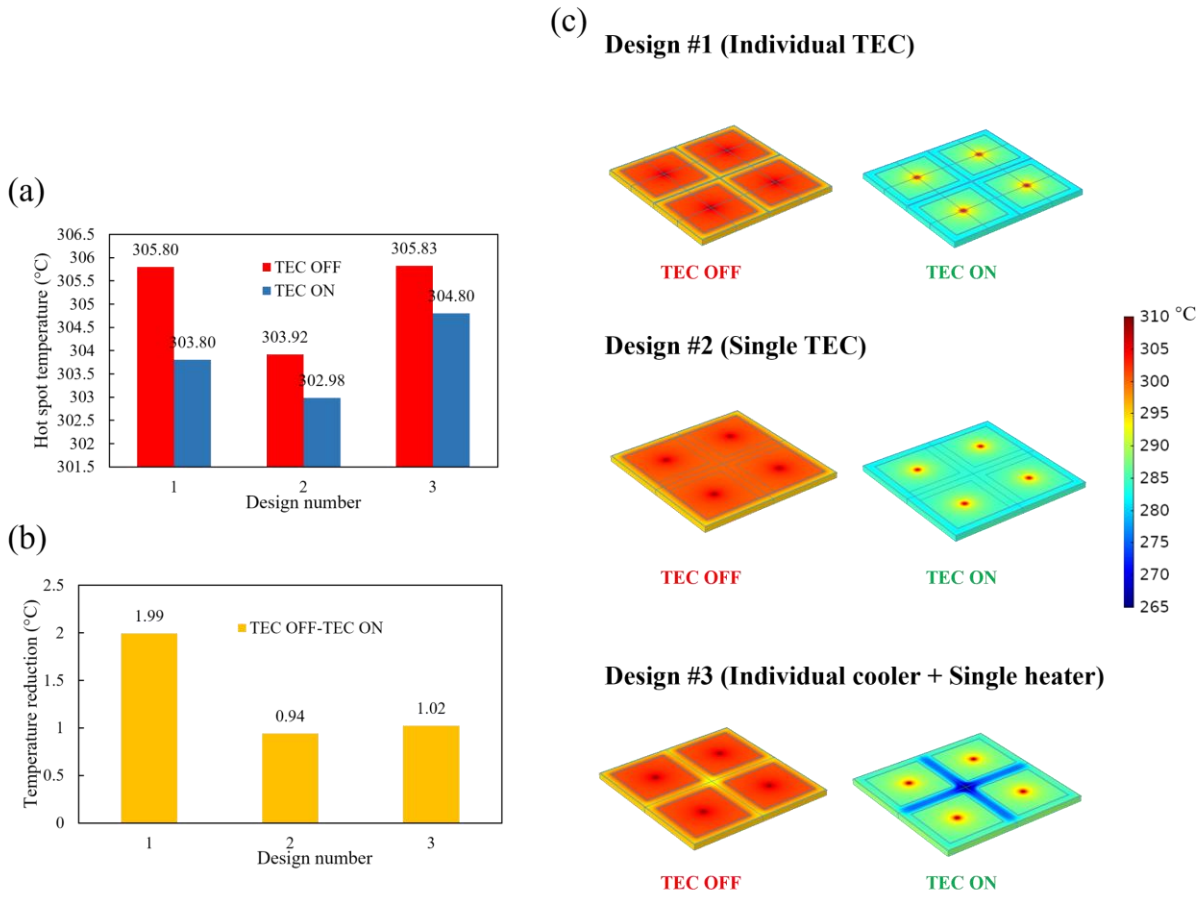

Figure 5-4. Results of steady-state reference simulation: (a) the steady-state hotspot temperature, (b) the hotspot temperature reduction, (c) the temperature distribution map.

In the reference simulation, design #1 has the greatest temperature reduction while design #2 has the lowest hotspot temperature. Although better TEC cooling performance can be found in design #1, more significant passive cooling occurs in design #2 in the absence of low thermal conductivity TE material inside the chip. On the other hand, design #3 achieves a moderate temperature reduction due to reduced heater area.

5.2.2 Effect of boundary area

The 2D parametric study (i.e., section 3.1.4 and 3.1.7) indicates that increasing the length of the TEC can lead to enhanced TEC cooling performance. Similarly, in the array design, this effect can be achieved by incorporating additional TEC boundary area. Figure 5-5 demonstrates the optimization results after implementing an extra TEC boundary (i.e., Δb) of 15, 30 and 50 µm. Figure 5-5 (b) shows that such increase in TEC boundary length provides enhanced temperature reduction for all array designs. Specifically, design #2 and #3 shows greater enhancement due to the presence of single external heater.

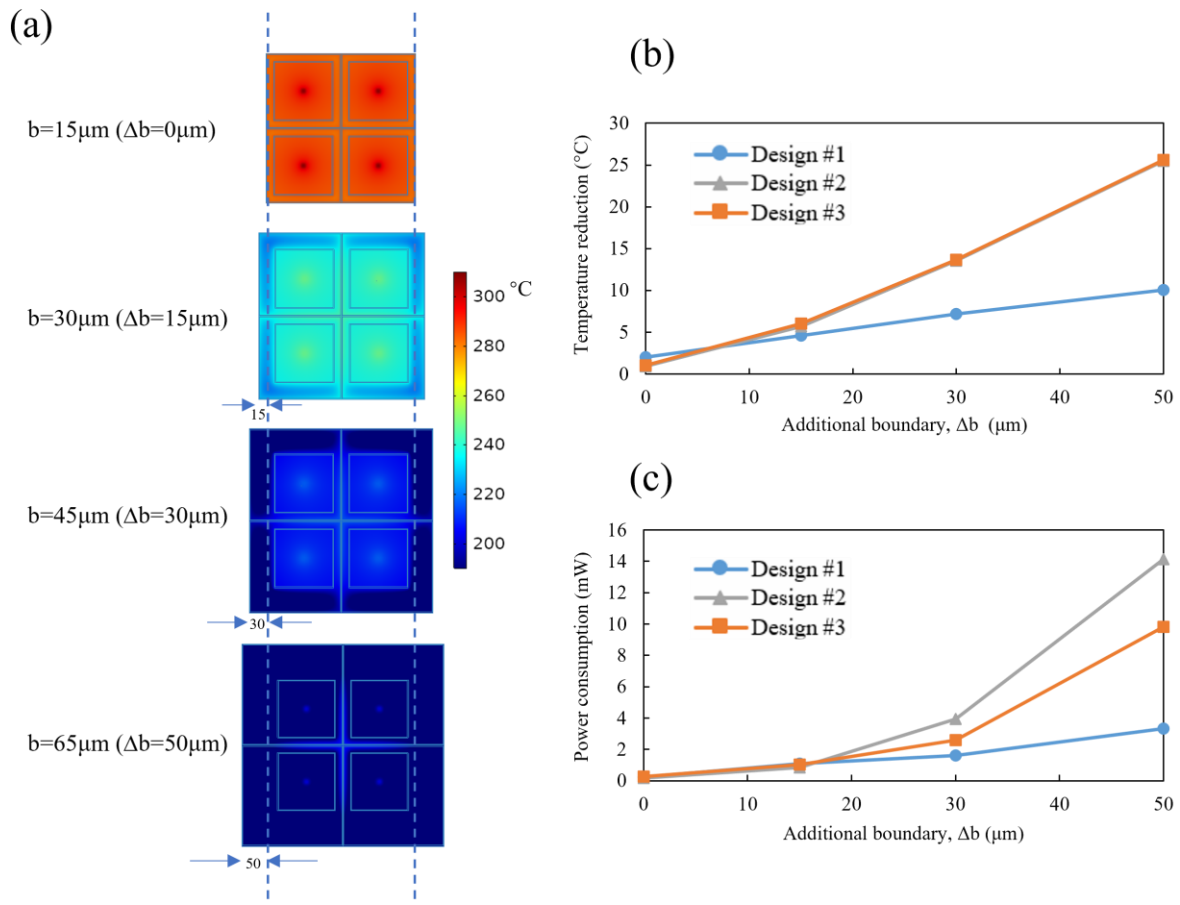

Figure 5-5. Optimization with increased TEC boundary lengths: (a) temperature distribution maps for different boundary lengths, (b) relationship between temperature reduction and additional boundary, and (c) power consumption as a function of additional boundary.

The hotspot temperature and the corresponding temperature reduction are shown in Figure 5-6. With additional boundary length, although design #2 exhibits the lowest TEC ON temperature, design #1 demonstrates the lowest TEC OFF temperature. Notably, expanding the boundary area can benefit both active TEC cooling and passive cooling. Consequently, when selecting the array designs, it is crucial to consider low TEC OFF temperature due to the additional power consumption associated with activating the TEC.

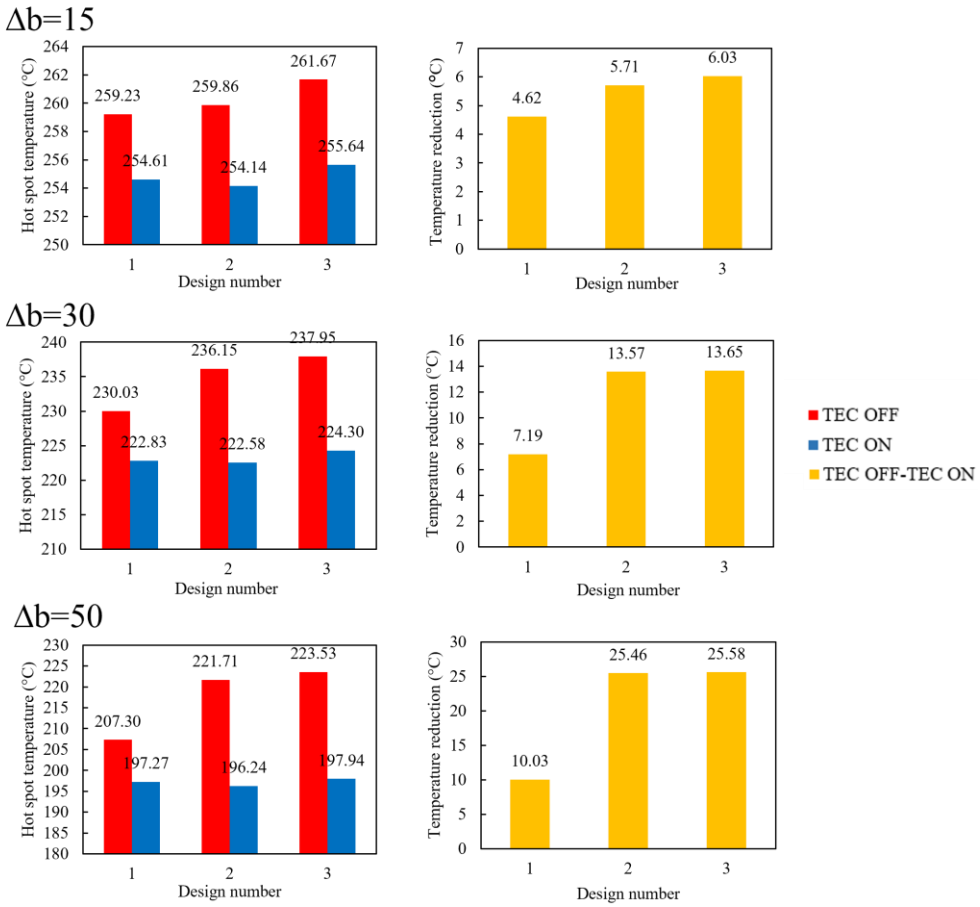

Figure 5-6. The hotspot temperature and the corresponding temperature reduction data in three array designs with increased TEC boundary lengths: (left) hotspot temperature and (right) temperature reduction.

5.2.3 Effect of internal distance

In the transistor-TEC array, the internal distance determines the gap length between each transistor-TEC unit. In design #1 and #3, this gap is filled with holey silicon, while in design #2, however, it consists of bulk silicon.

The temperature reduction is enhanced with an increase in gap distance for all three array designs, as depicted in Figure 5-7. However, the most significant enhancement is observed in design #1 due to the presence of individual TECs. With a larger gap, the heaters of the individual TECs can facilitate more efficient heat exchange with the external heat sink.

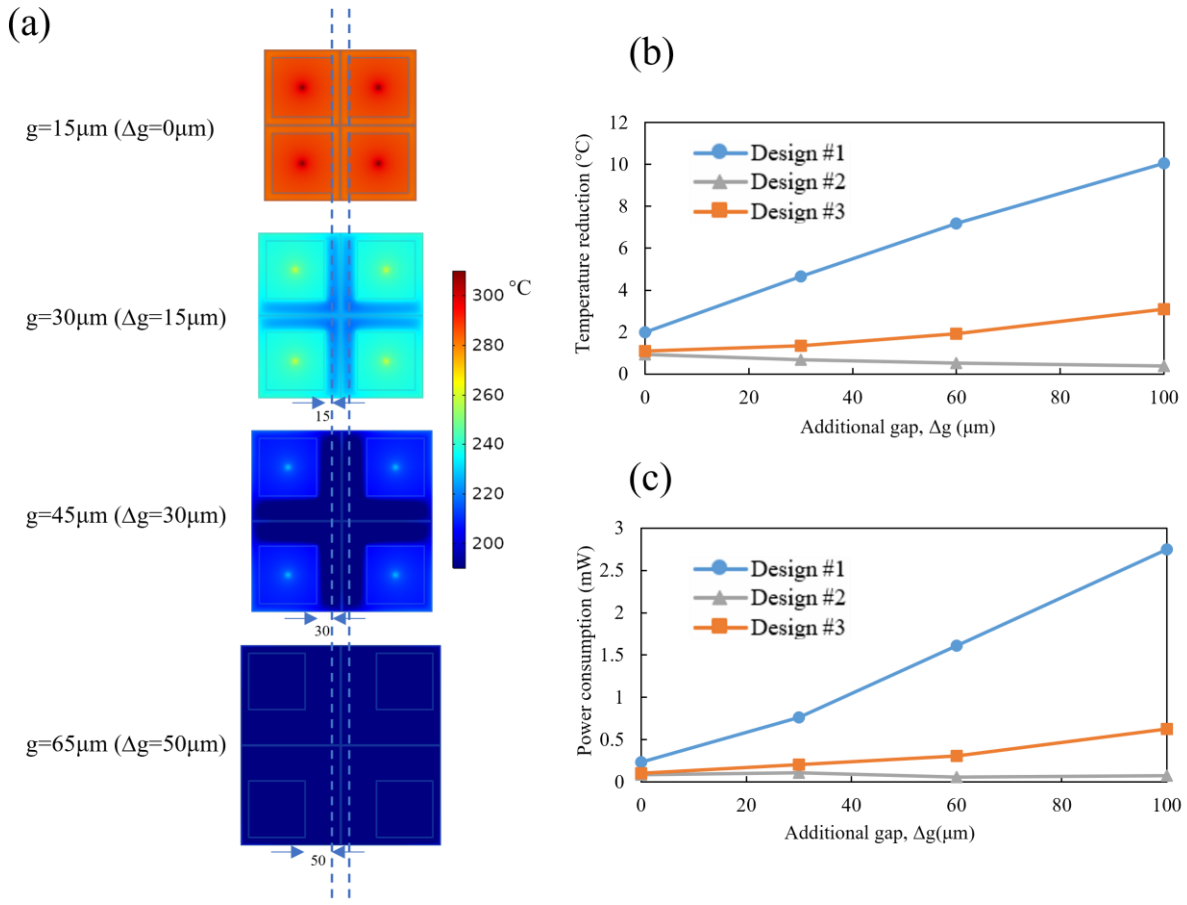

Figure 5-7. Optimization with increasing internal distance: (a) temperature distribution maps for different internal distance, (b) relationship between temperature reduction and additional gap distance, and (c) power consumption as a function of additional gap distance.

Figure 5-8 demonstrates the hotspot temperature and the corresponding temperature reduction data in three array designs with increasing internal distance. While design #1 shows more promising TEC cooling performance (i.e, a greater temperature reduction between TEC OFF and TEC ON states), design #2 demonstrates much lower TEC ON and OFF temperatures due to notable enhancements in passive cooling efficiency. Consequently, an optimal TEC array device should strike a balance between active TEC cooling and passive cooling.

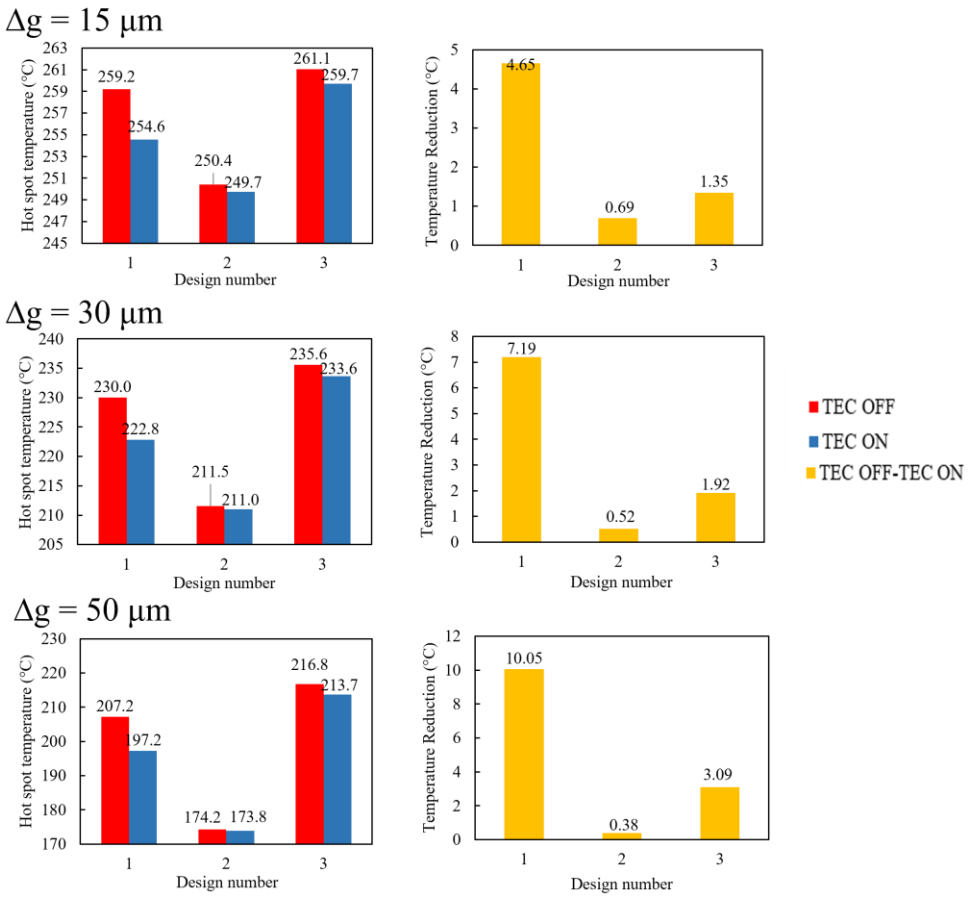

Figure 5-8. The hotspot temperature and the corresponding temperature reduction data in three array designs with increasing internal distance: (left) hotspot temperature and (right) temperature reduction.

5.3 Preliminary Study of Dynamic Hotspot

The previous array analysis assumes that all 4 hotspots in the 2 × 2 transistor matrix is activated simultaneously. However, in practical applications, the transistor matrix may undergo dynamic hotspot conditions where only a few of all the hotspots are turned on. Figure 5-9 demonstrates an example of the dynamic hotspot conditions in three array designs, where only one hotspot in the 2 × 2 transistor matrix is activated.

Due to the asymmetry of the hotspot conditions, for design #1 and #3, the TEC coolers can have multiple options of activation: (a) activate the corresponding cooler, (b) activate the corresponding cooler and the adjacent coolers, (c) activate the adjacent coolers and (d) activate all the coolers. Moreover, these options can be even more complicated if different TEC currents are applied in different coolers. For the preliminary study, we assume the only option for the TEC activation is to activate the corresponding cooler. Table 5-2 specifies the parameters used in the array simulation with dynamic hotspot conditions.

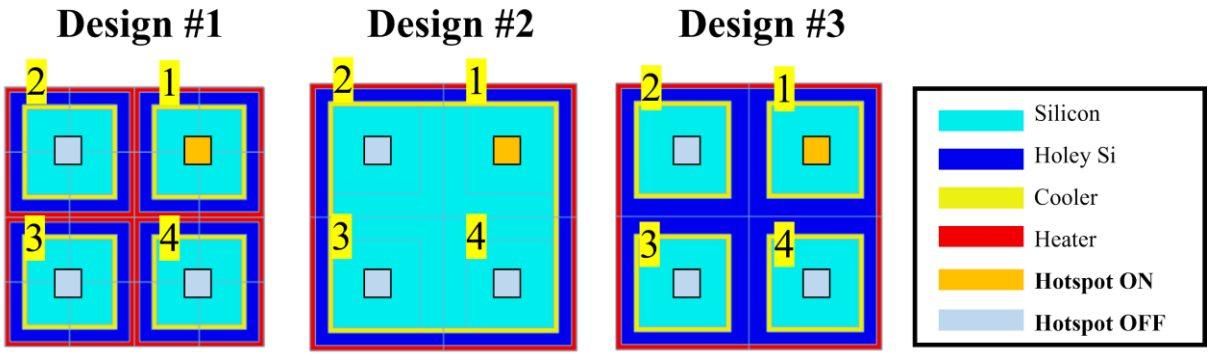

Figure 5-9. Schematics of dynamic hotspot in three array designs. Only one hotspot in the 2 × 2 transistor matrix is activated.

Geometrical parameters

Parameters	Name	Value
a	Hotspot distance	50 μm
b	TEC length	15 μm
c	Length (x) of cooler	1.5 μm
d	Length (x) of ground	1.5 μm
e	Thickness (z) of chip	10 μm
f	Distance between hotspot and top surface	1 μm
L_{HS}	Length (x) of hotspot	4 μm
T_{HS}	Thickness (z) of hotspot	1 μm

Material properties & boundary conditions

Parameters	Name	Value
h_{eff}	Effective convection coefficient	5000 W/m^{2}K
Q	Hotspot heat flux	35000 W/cm^2
T_∞	Ambient temperature	20 °C
$k_{HS,x}$	In-plane thermal conductivity of holey silicon	1 W/mK
$k_{HS,y}$	Cross-plane thermal conductivity of holey silicon	13 W/mK
k_{Si}	Thermal conductivity of bulk silicon	65 W/mK
k_{Al}	Thermal conductivity of aluminum	237 W/mK
S_{Si}	Seebeck coefficient of (holey) silicon	440 μV/K

Table 5-2. Specific parameters used in the array simulation with dynamic hotspot conditions.

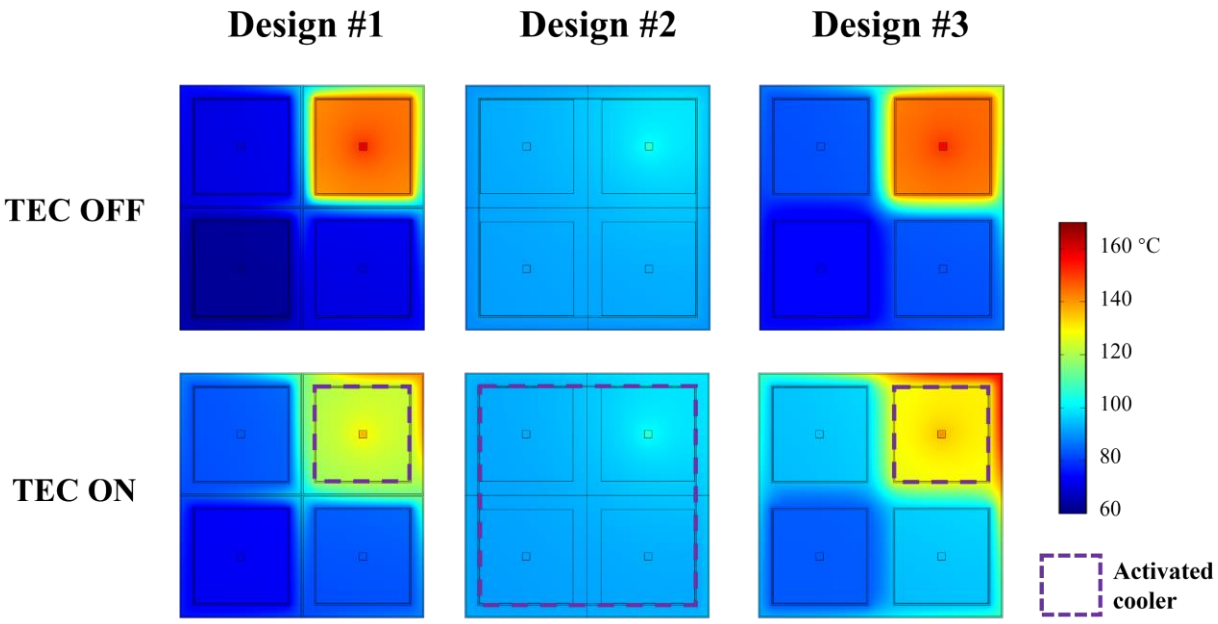

Figure 5-10. Temperature distribution maps of three array designs under dynamic hotspot conditions.

The temperature distribution maps of three array designs are shown in Figure 5-10 under dynamic hotspot conditions. Only the hotspot #1 (top-right corner) is activated, together with the corresponding TEC cooler. Figure 5-11 shows the hotspot temperature as a function of the TEC current. Four hotspot locations in the 2 × 2 matrix (labeled as HS #1 to #4) are evaluated.

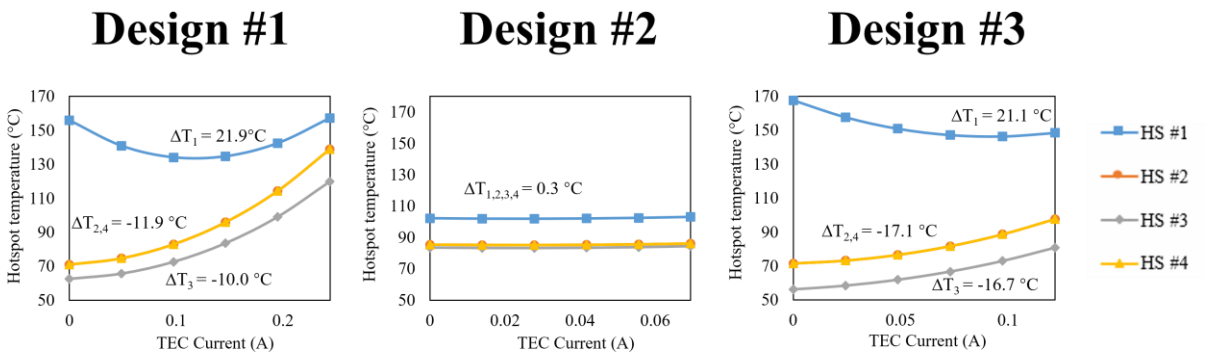

Figure 5-11. Hotspot temperature as a function of the TEC current in four hotspot locations.

Among three array designs, design #1 and #3 exhibit significant temperature changes at hotspot #1. Specifically, with optimal TEC current, a maximum temperature reduction of 21.9 °C and 21.1°C is achieved at the activated hotspot in design #1 and #3, respectively. However, this substantial reduction also leads to a temperature increase at hotspot #2, #3 and #4 due to the asymmetric boundary conditions that transport heat from the activated heater to the nearby inactivated regions.

While the temperature reduction is significantly greater for design #1 and #3, the absolute temperatures in both ON and OFF states are notably lower for design #2. Due to the absence of holey silicon in its internal region, design #2 exhibits more pronounced passive cooling. Consequently, further optimization is required to provide high-performance TEC cooling under dynamic hotspot conditions.

Chapter 6

Conclusion and Outlooks

While the state-of-the-art commercial microelectronic device has achieved a technology node as small as 5nm, and even below 5nm, most thermal management techniques are limited to the macroscale level. The innovative concept of designing a microscale holey silicon-based lateral TEC aims to address the increasingly demanding yet delicate thermal management challenges in the microelectronic system. Holey silicon, a nanostructured silicon with excellent thermoelectric properties and compatibility with the conventional microfabrication process, enables direct integration of the TEC into the architecture, creating an all-in-one cooling system.

In this thesis, a preliminary study of the transistor-TEC system is performed using the COMSOL Multiphysics platform. With the increasing understanding of the TEC design, our investigation and optimization go from 2D to 3D, from steady state to transient state, and from single TEC to TEC array. Many rules of thumb regarding the TEC optimization are found during the parametric study in section 3.1, 3.2, 4.2 and 5.2. In general, the hotspot-cooler distance as well as the TEC length act as the most important geometrical parameters in the design optimization of the transistor-TEC system. Besides, the ambient temperature, the efficiency of heat sink (i.e., the heat convection coefficient) as well as the heating conditions play a crucial role in determining the TEC cooling performance.

Besides maximizing the TEC cooling performance (i.e., the hotspot temperature reduction), the design of transistor-TEC system should consider other factors including the space confinements, passive cooling and the economical factor (i.e., power consumption). Generally speaking, a high TEC cooling performance can only be obtained at the expense of a large lateral TEC area and a high TEC power. These trade-off effects should be seriously reviewed when designing a space- and power-efficient microelectronic device. Moreover, an efficient TEC also results in degradation of passive cooling due to the introduction of low thermal conductivity thermoelectric material. Detailed analysis regarding the trade-off effects among the abovementioned factors can be found in section 3.2 and 5.2.

While the steady-state TEC shows promising cooling performance under stationary heating conditions, it is more applicable and effective to implement the transient TEC cooling. To investigate the transient TEC effect, we utilize a transient TEC current pulse to cool down both a constant heat flux (section 4.1) and a transient heat pulse (section 4.2). For both scenarios, the supercooling effect is observed during the simulation, where a temporal mismatch of the Peltier effect and the volumetric Joule effect results in a much greater temperature reduction, followed by a delayed temperature overshoot. For some situations, this temperature overshoot may heat up the device instead. At this point, an optimal TEC current pulse helps to alleviate the side effect and maximize the cooling. More detailed discussion regarding the TEC current pulse optimization, including amplitude, temporal location, duration and shape is performed in section 4.2.

Nowadays, the integration of high-power-density transistor arrays have become a common practice, which enlightens us to explore the possibilities of TEC array design. In chapter 5, we perform the simulation of transistor-TEC array system. Three different array designs are demonstrated for comparison and the preliminary optimization results are performed. In brief, the

individual TEC design (i.e., design #1) has the best TEC cooling performance while the single TEC design (i.e., design #2) provides the lowest TEC OFF temperature. Detailed analysis regarding uniform hotspot condition and dynamic hotspot condition can be found separately in section 5.2 and 5.3.

In conclusion, the holey silicon-based lateral TEC has demonstrated promising TEC cooling performance for the microelectronic devices. Future efforts will be focusing on the TEC device demonstration, verification of the simulation results, and machine learning-based optimal cooling. Actual fabrication will be carried out in the cleanroom using lithography, dry etching and metal deposition. Further electrical and thermal characterization will be performed using the actual TEC device.

References

[1] I. CIRCUITS, "CRAMMING MORE," *Understanding Moore's Law: Four Decades of Innovation*, p. 55, 2006.

[2] Todd Wilder, "Apple introduces M2 Ultra," *Apple*, Jun. 05, 2023.

[3] S. Kaxiras and M. Martonosi, *Computer architecture techniques for power-efficiency*. Springer Nature, 2022.

[4] Z. Lun, G. Du, J. Qin, Y. Wang, J. Wang, and X. Liu, "Investigation of self-heating effect in SOI-LDMOS by device simulation," in *ICSICT 2012 - 2012 IEEE 11th International Conference on Solid-State and Integrated Circuit Technology, Proceedings*, 2012. doi: 10.1109/ICSICT.2012.6467622.

[5] H. T. Lim, F. Udrea, D. M. Garner, and W. I. Milne, "Modelling of self-heating e€ect in thin SOI and Partial SOI LDMOS power devices."

[6] M. Pfost, J. Joos, and M. Stecher, "Measurement and simulation of self-heating in DMOS transistors up to very high temperatures," in *Proceedings of the International Symposium on Power Semiconductor Devices and ICs*, Institute of Electrical and Electronics Engineers Inc., 2008, pp. 209–212. doi: 10.1109/ISPSD.2008.4538935.

[7] A. Gupta, C. Gupta, R. A. Vega, T. B. Hook, and A. Dixit, "Reliability Modeling and Analysis of Hot-Carrier Degradation in Multiple-Fin SOI n-Channel FinFETs With Self-Heating," *IEEE Trans Electron Devices*, vol. 66, no. 5, pp. 2075–2080, May 2019, doi: 10.1109/TED.2019.2905053.

[8] A. N. Tallarico *et al.*, "Hot-Carrier Degradation in Power LDMOS: Selective LOCOS-Versus STI-Based Architecture," *IEEE Journal of the Electron Devices Society*, vol. 6, no. 1, pp. 219–226, 2018, doi: 10.1109/JEDS.2018.2792539.

[9] D. Dibra, M. Stecher, S. Decker, A. Lindemann, J. Lutz, and C. Kadow, "On the origin of thermal runaway in a trench power MOSFET," *IEEE Trans Electron Devices*, vol. 58, no. 10, pp. 3477–3484, Oct. 2011, doi: 10.1109/TED.2011.2160867.

[10] D. Simon, C. Boianceanu, G. De Mey, V. Topa, and A. Spitzer, "Reliability Analysis for Power Devices Which Undergo Fast Thermal Cycling," *IEEE Transactions on Device and Materials Reliability*, vol. 16, no. 3, pp. 336–344, Sep. 2016, doi: 10.1109/TDMR.2016.2583663.

[11] P. Rodgers, V. Eveloy, and M. G. Pecht, "Limits of Air-Cooling: Status and Challenges."

[12] M. Saini and R. L. Webb, "HEAT REJECTION LIMITS OF AIR COOLED PLANE FIN HEAT SINKS FOR COMPUTER COOLING."

[13] G. Xu, B. Guenin, and M. Vogel, "Extension of air cooling for high power processors," in *The Ninth Intersociety Conference on Thermal and Thermomechanical Phenomena In Electronic Systems (IEEE Cat. No. 04CH37543)*, IEEE, 2004, pp. 186–193.

[14] A. Habibi Khalaj and S. K. Halgamuge, "A Review on efficient thermal management of air- and liquid-cooled data centers: From chip to the cooling system," *Applied Energy*, vol. 205. Elsevier Ltd, pp. 1165–1188, Nov. 01, 2017. doi: 10.1016/j.apenergy.2017.08.037.

[15] Y. J. Lee, P. K. Singh, and P. S. Lee, "Fluid flow and heat transfer investigations on enhanced microchannel heat sink using oblique fins with parametric study," *Int J Heat Mass Transf*, vol. 81, pp. 325–336, 2015.

[16] C. H. Hoang *et al.*, "A Review of Recent Developments in Pumped Two-Phase Cooling Technologies for Electronic Devices," *IEEE Trans Compon Packaging Manuf Technol*, 2021, doi: 10.1109/TCPMT.2021.3117572.

[17] K. P. Drummond *et al.*, "A hierarchical manifold microchannel heat sink array for high-heat-flux two-phase cooling of electronics," *Int J Heat Mass Transf*, vol. 117, pp. 319–330, 2018.

[18] C. Kaviarasu and D. Prakash, "Review on Phase Change Materials with Nanoparticle in Engineering Applications.," *Journal of Engineering Science & Technology Review*, vol. 9, no. 4, 2016.

[19] R. D. Pathumudy and K. N. Prabhu, "Thermal interface materials for cooling microelectronic systems: present status and future challenges," *Journal of Materials Science: Materials in Electronics*, vol. 32, no. 9. Springer, pp. 11339–11366, May 01, 2021. doi: 10.1007/s10854-021-05635-w.

[20] H. Lv, X. D. Wang, T. H. Wang, and J. H. Meng, "Optimal pulse current shape for transient supercooling of thermoelectric cooler," *Energy*, vol. 83, pp. 788–796, Apr. 2015, doi: 10.1016/j.energy.2015.02.092.

[21] L. E. Bell, “Cooling, heating, generating power, and recovering waste heat with thermoelectric systems,” *Science (1979)*, vol. 321, no. 5895, pp. 1457–1461, 2008.

[22] B. C. Sales, “Smaller is cooler,” *Science (1979)*, vol. 295, no. 5558, pp. 1248–1249, 2002.

[23] D. M. Rowe, *CRC handbook of thermoelectrics*. CRC press, 2018.

[24] G. Bulman *et al.*, “Superlattice-based thin-film thermoelectric modules with high cooling fluxes,” *Nat Commun*, vol. 7, Jan. 2016, doi: 10.1038/ncomms10302.

[25] G. E. Bulman, E. Siivola, R. Wiitala, R. Venkatasubramanian, M. Acree, and N. Ritz, “Three-stage thin-film superlattice thermoelectric multistage microcoolers with a Δt max of 102 K,” in *Journal of Electronic Materials*, Jul. 2009, pp. 1510–1515. doi: 10.1007/s11664-009-0742-2.

[26] G. E. Bulman, E. Siivola, B. Shen, and R. Venkatasubramanian, “Large external ΔT and cooling power densities in thin-film Bi 2Te3-superlattice thermoelectric cooling devices,” *Appl Phys Lett*, vol. 89, no. 12, 2006, doi: 10.1063/1.2353805.

[27] R. Yang, G. Chen, A. R. Kumar, G. J. Snyder, and J. P. Fleurial, “Transient cooling of thermoelectric coolers and its applications for microdevices,” *Energy Convers Manag*, vol. 46, no. 9–10, pp. 1407–1421, Jun. 2005, doi: 10.1016/j.enconman.2004.07.004.

[28] Z. Ren, J. C. Kim, and J. Lee, “Transient Cooling and Heating Effects in Holey Silicon-Based Lateral Thermoelectric Devices for Hot Spot Thermal Management,” *IEEE Trans Compon Packaging Manuf Technol*, vol. 11, no. 8, pp. 1214–1222, Aug. 2021, doi: 10.1109/TCPMT.2021.3081375.

[29] G. J. Snyder, J. P. Fleurial, T. Caillat, R. Yang, and G. Chen, "Supercooling of Peltier cooler using a current pulse," *J Appl Phys*, vol. 92, no. 3, pp. 1564–1569, Aug. 2002, doi: 10.1063/1.1489713.

[30] A. J. Piggott and J. S. Allen, "Peltier Supercooling with Isosceles Current Pulses: A Response Surface Perspective," *ECS Journal of Solid State Science and Technology*, vol. 6, no. 3, pp. N3045–N3054, 2017, doi: 10.1149/2.0061703jss.

[31] A. Shakouri and Y. Zhang, "On-chip solid-state cooling for integrated circuits using thin-film microrefrigerators," *IEEE Transactions on Components and Packaging Technologies*, vol. 28, no. 1, pp. 65–69, Mar. 2005, doi: 10.1109/TCAPT.2005.843219.

[32] P. Wang and A. Bar-Cohen, "On-chip hot spot cooling using silicon thermoelectric microcoolers," *J Appl Phys*, vol. 102, no. 3, 2007, doi: 10.1063/1.2761839.

[33] I. Chowdhury *et al.*, "On-chip cooling by superlattice-based thin-film thermoelectrics," *Nat Nanotechnol*, vol. 4, no. 4, pp. 235–238, 2009, doi: 10.1038/nnano.2008.417.

[34] Y. Su, J. Lu, and B. Huang, "Free-standing planar thin-film thermoelectric microrefrigerators and the effects of thermal and electrical contact resistances," *Int J Heat Mass Transf*, vol. 117, pp. 436–446, Feb. 2018, doi: 10.1016/j.ijheatmasstransfer.2017.10.023.

[35] Y. Su, J. Lu, D. Villaroman, D. Li, and B. Huang, "Free-standing planar thermoelectric microrefrigerators based on nano-grained SiGe thin films for on-chip refrigeration," *Nano Energy*, vol. 48, pp. 202–210, Jun. 2018, doi: 10.1016/j.nanoen.2018.03.054.

[36] X. Nie *et al.*, "Numerical Simulation and Structural Optimization of Multi-Stage Planar Thermoelectric Coolers," *physica status solidi (a)*, vol. 217, no. 22, p. 2000248, 2020.

[37] D. Zhang *et al.*, "Multifunctional superelastic graphene-based thermoelectric sponges for wearable and thermal management devices," *Nano Lett*, vol. 22, no. 8, pp. 3417–3424, 2022.

[38] T. C. Harman, P. J. Taylor, M. P. Walsh, and B. E. Laforge, "Quantum Dot Superlattice Thermoelectric Materials and Devices." [Online]. Available: https://www.science.org

[39] X. Fan *et al.*, "SiGeC/Si superlattice microcoolers," *Appl Phys Lett*, vol. 78, no. 11, pp. 1580–1582, Mar. 2001, doi: 10.1063/1.1356455.

[40] B. Yang, P. Wang, and A. Bar-Cohen, "Mini-contact enhanced thermoelectric cooling of hot spots in high power devices," *IEEE Transactions on Components and Packaging Technologies*, vol. 30, no. 3, pp. 432–438, 2007, doi: 10.1109/TCAPT.2007.898744.

[41] P. Wang, B. Yang, and A. Bar-Cohen, "Mini-contact enhanced thermoelectric coolers for on-chip hot spot cooling," *Heat Transfer Engineering*, vol. 30, no. 9, pp. 736–743, Aug. 2009, doi: 10.1080/01457630802678391.

[42] H. Kattan, S. W. Chung, J. Henkel, and H. Amrouch, "On-Demand Mobile CPU Cooling with Thin-Film Thermoelectric Array," *IEEE Micro*, vol. 41, no. 4, pp. 67–73, Jul. 2021, doi: 10.1109/MM.2021.3061335.

[43] H. Alam and S. Ramakrishna, "A review on the enhancement of figure of merit from bulk to nano-thermoelectric materials," *Nano Energy*, vol. 2, no. 2. pp. 190–212, Mar. 2013. doi: 10.1016/j.nanoen.2012.10.005.

[44] W. Zhao, J. Ding, Y. Zou, C. A. Di, and D. Zhu, "Chemical doping of organic semiconductors for thermoelectric applications," *Chemical Society Reviews*, vol. 49, no. 20. Royal Society of Chemistry, pp. 7210–7228, Oct. 21, 2020. doi: 10.1039/d0cs00204f.

[45] G. J. Snyder and E. S. Toberer, "Complex thermoelectric materials," *Nat Mater*, vol. 7, no. 2, pp. 105–114, 2008.

[46] J. Ma, D. Gelda, K. V. Valavala, and S. Sinha, "Peak thermoelectric power factor of holey silicon films," *J Appl Phys*, vol. 128, no. 11, Sep. 2020, doi: 10.1063/5.0010254.

[47] J. Tang *et al.*, "Holey silicon as an efficient thermoelectric material," *Nano Lett*, vol. 10, no. 10, pp. 4279–4283, Oct. 2010, doi: 10.1021/nl102931z.

[48] J. Lee, J. Lim, and P. Yang, "Ballistic phonon transport in holey silicon," *Nano Lett*, vol. 15, no. 5, pp. 3273–3279, May 2015, doi: 10.1021/acs.nanolett.5b00495.

[49] J. Lim *et al.*, "Simultaneous thermoelectric property measurement and incoherent phonon transport in holey silicon," *ACS Nano*, vol. 10, no. 1, pp. 124–132, Jan. 2016, doi: 10.1021/acsnano.5b05385.

[50] N. Liu, T. Zhu, M. G. Rosul, J. Peters, J. E. Bowers, and M. Zebarjadi, "Thermoelectric properties of holey silicon at elevated temperatures," *Materials Today Physics*, vol. 14, Aug. 2020, doi: 10.1016/j.mtphys.2020.100224.

[51] Z. Ren and J. Lee, "Thermal conductivity anisotropy in holey silicon nanostructures and its impact on thermoelectric cooling," *Nanotechnology*, vol. 29, no. 4, Jan. 2018, doi: 10.1088/1361-6528/aa9f07.

[52] R. Chen, A. I. Hochbaum, P. Murphy, J. Moore, P. Yang, and A. Majumdar, "Thermal conductance of thin silicon nanowires," *Phys Rev Lett*, vol. 101, no. 10, p. 105501, 2008.

[53] M. G. Holland, "Analysis of lattice thermal conductivity," *Physical review*, vol. 132, no. 6, p. 2461, 1963.

[54] M. Asheghi, M. N. Touzelbaev, K. E. Goodson, Y. K. Leung, and S. S. Wong, "Temperature-dependent thermal conductivity of single-crystal silicon layers in SOI substrates," 1998.

[55] W. Liu and M. Asheghi, "Thermal conduction in ultra-thin pure and doped single crystal silicon layers at high temperatures," in *Heat Transfer Summer Conference*, 2005, pp. 821–827.

[56] E. H. Sondheimer, "The mean free path of electrons in metals," *Adv Phys*, vol. 50, no. 6, pp. 499–537, 2001.

[57] G. Chen, "Thermal conductivity and ballistic-phonon transport in the cross-plane direction of superlattices," *Phys Rev B*, vol. 57, no. 23, p. 14958, 1998.

[58] Y. S. Ju and K. E. Goodson, "Phonon scattering in silicon films with thickness of order 100 nm," *Appl Phys Lett*, vol. 74, no. 20, pp. 3005–3007, 1999.

[59] B. Qiu *et al.*, "First-Principles Simulation of Electron Mean-Free-Path Spectra and Thermoelectric Properties in Silicon."

[60] D. Neamen, *Semiconductor physics and devices*. McGraw-Hill, Inc., 2002.

[61] Aria Hosseini, "ThermoElectric," Apr. 05, 2022.

[62] P. Wang, A. Bar-Cohen, B. Yang, G. L. Solbrekken, and A. Shakouri, "Analytical modeling of silicon thermoelectric microcooler," *J Appl Phys*, vol. 100, no. 1, 2006, doi: 10.1063/1.2211328.

[63] T. H. R. F. Gmaar, G. W. And, and L. Hur, "Seebeck Effect in Silicon," 1955.

[64] F. H. Horn, "Densitometric and Electrical Investigation of Boron in Silicon," *Physical Review*, vol. 97, no. 6, pp. 1521–1525, Mar. 1955, doi: 10.1103/PhysRev.97.1521.

[65] J. L. Braun *et al.*, "Size effects on the thermal conductivity of amorphous silicon thin films," *Phys Rev B*, vol. 93, no. 14, Apr. 2016, doi: 10.1103/PhysRevB.93.140201.

[66] W. Liu and M. Asheghi, "Thermal conductivity measurements of ultra-thin single crystal silicon layers," *J Heat Transfer*, vol. 128, no. 1, pp. 75–83, Jan. 2006, doi: 10.1115/1.2130403.

[67] H. Lv, X. D. Wang, T. H. Wang, and J. H. Meng, "Optimal pulse current shape for transient supercooling of thermoelectric cooler," *Energy*, vol. 83, pp. 788–796, Apr. 2015, doi: 10.1016/j.energy.2015.02.092.

[68] C. Selvam, S. Manikandan, S. C. Kaushik, R. Lamba, and S. Harish, "Transient performance of a Peltier super cooler under varied electric pulse conditions with phase change material," *Energy Convers Manag*, vol. 198, Oct. 2019, doi: 10.1016/j.enconman.2019.111822.

[69] M. A. Belaïd, K. Ketata, K. Mourgues, M. Gares, M. Masmoudi, and J. Marcon, "Reliability study of power RF LDMOS device under thermal stress," *Microelectronics J*, vol. 38, no. 2 SPEC. ISS., pp. 164–170, 2007, doi: 10.1016/j.mejo.2006.08.004.

[70] A. Raman, D. G. Walker, and T. S. Fisher, “Simulation of nonequilibrium thermal effects in power LDMOS transistors,” *Solid State Electron*, vol. 47, no. 8, pp. 1265–1273, Aug. 2003, doi: 10.1016/S0038-1101(03)00066-2.

[71] T. Bucher and H. Amrouch, “Modeling TPU Thermal Maps Under Superlattice Thermoelectric Cooling,” *IEEE Access*, vol. 10, pp. 21970–21978, 2022, doi: 10.1109/ACCESS.2022.3152195.

[72] H. Amrouch, G. Zervakis, S. Salamin, H. Kattan, I. Anagnostopoulos, and J. Henkel, “NPU Thermal Management,” in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 3842–3855. doi: 10.1109/TCAD.2020.3012753.

[73] Karl Rupp, “50 Years of Microprocessor Trend Data,” *Github*, Feb. 22, 2022.

[74] J.-F. Li, W.-S. Liu, L.-D. Zhao, and M. Zhou, “High-performance nanostructured thermoelectric materials,” *NPG Asia Mater*, vol. 2, no. 4, pp. 152–158, 2010.

[75] K. Maize *et al.*, “Thermal characterization of high power transistor arrays,” in *2009 25th Annual IEEE Semiconductor Thermal Measurement and Management Symposium*, IEEE, 2009, pp. 50–54.