COMPACT MODELING METHODOLOGY DEVELOPMENT FOR THERMO-MECHANICAL ASSESSMENT IN HIGH-END MOBILE APPLICATIONS – PLANAR AND 3D TSV PACKAGES

by

FAHAD MIRZA

Presented to the Faculty of the Mechanical Engineering of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2014

ACKNOWLEDGEMENTS

I thank Dr. Agonafer for advising me through my PhD. and guiding me to this successful juncture. A significant portion of this work was performed during an internship at GLOBALFOUNDRIES, Malta, NY. I sincerely thank Mr. Premachandran C S (*SMTS- QRA*), Mr. Chirag Shah (Ex Principal Engineer, QRA) and Dr. Natarajan Mahadeva Iyer (Director-QRA), to give me an opportunity and guide me through the internship that enabled it to be a worthwhile venture for GLOBALFOUNDRIES and me.

I also want to thank all my committee members, Dr. Seiichi Nomura, Dr. Zeynup Celik-Butler, Dr. Abdolhossein Haji-Sheikh, and Dr. Natarajan Mahadeva Iyer for being on my committee and giving their valuable inputs and time that enabled me to successfully complete this work.

I cannot miss to acknowledge my team of graduate students, Thiagarajan Raman (currently at GLOBALFOUNDRIES), Zaeem Baig (currently at Zimmer Inc.), Hardik Parekh (PVC Fluids) and Hassaan Ahmad Khan, Tejas Shetty, Pavan Rajmane, and Abhishek Deshpande (current MS students in EMNSPC) who helped me technically and emotionally through this long process. I also had the opportunity to mentor them through their graduate degree and co-publish with them.

November 17, 2014

ABSTRACT

COMPACT MODELING METHODOLOGY DEVELOPMENT FOR THERMO-MECHANICAL ASSESSMENT IN HIGH-END MOBILE APPLICATIONS – PLANAR AND 3D TSV PACKAGES

Fahad Mirza, PhD

The University of Texas at Arlington, 2014

Supervising Professor: Dereje Agonafer

For the last few decades, Moore's law has been driving the microelectronics industry to achieve high performance with small form-factors at a reasonable cost. The continued miniaturization of the transistors has resulted in unparalleled growth in the electronics industry. Further performance increment via size scaling; however, will not be cost-effective and difficult to manufacture. To achieve further performance enhancement at a reasonable cost, advanced packaging is being leveraged. Advanced packaging technologies such as System-on-chip (SOCs), system-in-package (SiPs), and 3-D TSV ICs provide significant power and performance enhancement without having the need to migrate to the higher technology node. As planar device miniaturization continues to its ultimate limits, the complexity of circuit interconnections for 2-D devices becomes a limitation for performance and drives up power dissipation. 3-D technology would enable extremely dense solid-state memory to be arrayed within a few microns of the processing elements, which reduces access times. It is clear that these advanced packaging

technologies are advantageous from cost and performance standpoint, and are therefore shaping up as an integral part of the mainstream consumer electronics, especially handheld applications. However, the development and sustainability of such 'big things" requires immense numerical and experimental testing in order to have a reliable device. Reliability and performance of such systems holds the key. Mechanical reliability needs to be evaluated and failures need to be mitigated for smooth and effective transition to the market. There is a need to develop sophisticated and compact numerical and experimental methodologies for efficient mechanical and performance analysis.

In this work, a novel compact modeling methodology has been developed to analyze the multi-level interconnects damage during process and product qualification for advanced high-end mobile application packages. The proposed technique has been rigorously validated against the full array models and literature and a correlation model to determine the modified ANAND's viscoplastic constants for the effective solder block compact model (for SAC 305) has been formulated. The proposed methodology has resulted in a significant computational time reduction during technology development leading to faster lead times.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii

ABSTRACT iv

LIST OF ILLUSTRATIONS ix

LIST OF TABLES xiv

Chapter 1 INTRODUCTION AND OBJECTIVE 1

1.1 Role of Packaging in Micro-Electronics 1

1.2 Motivation/Literature Review 2

Chapter 2 PLANAR PACKAGES – COMPACT MODELING METHODOLOGY DEVELOPMENT TO STUDY CHIP-PACKAGE INTERACTION IN FLIP CHIP PACKAGES 12

2.1 Chip-Package-Interaction (CPI) Prediction – Why is it a concern? 12

2.2 Literature Review 15

2.3 CPI Assessment on BEoL/fBEoL stack of large I/O chip 15

2.4 Compact 3D FEA Modeling 16

2.5 Linear property evaluation for compact modeling 19

2.6 Non-linear property evaluation for compact modeling 22

2.7 Results and Discussions 23

2.8 Acknowledgement 31

Chapter 3 PLANAR PACKAGES - SCALABILITY OF THE PROPOSED COMPACT MODELING METHODOLOGY – 2nd LEVEL BGA DAMAGE ASSESSMENT 32

3.1 Introduction 32

3.2 Literature Review 33
3.3 Compact Modeling and Methodology 35
3.4 Material Properties and Dimensions 38
3.5 Design of Experiments 41
3.6 Damage Prediction in BGAs (concept of the plastic work) 43
3.6.1 Energy Based Model 44
3.7 Results/Discussion 47
3.7.1 Linear effective block 48
3.7.2 Validation of SAC 305 Compact Model (non-linear block) 54
3.7.3 Qualitative validation of the compact model with literature (effect of mold compound elastic modulus on the corner joint response) 72
3.7.4 Validation of the compact modeling technique for different loading conditions (Accelerated Thermal Cycling - ATC) 78
3.8 Validation of the compact modeling technique across different solder alloys – Sn63Pb37 86
Chapter 4 3D TSV PACKAGE - COMPACT MODELING METHODOLOGY DEVELOPMENT FOR CPI ANALYSIS 91
4.1 3-D: Merits and Challenges 91
4.2 Introduction/ Literature Review 93
4.3 Model Description 101
4.3.1 Material Properties and Dimensions 105
4.3.2 Compact Model Development 106

4.4 Results and Discussion 108
4.4.1 Validation of Compact Modeling with Full Array model 108
4.4.2 Effect of TSV Joule Heating on Device Performance 115
Chapter 5 SUMMARY AND CONCLUSION 123
Bibliography 125
Biographical Statement 135

LIST OF ILLUSTRATIONS

Figure 1-1 Role of Packaging 1

Figure 1-2 Virtuous Cycle and constant component scaling [9] 3

Figure 1-3 Percent change in cost/gate [1] 4

Figure 1-4 Wafer cost increment with technology node [1] 6

Figure 1-5 “More Moore" and “More-Than-Moore” (MtM)” [9] 6

Figure 1-6 SiP/3D and SoC comparison for cost/function/time with system complexity .. 7

Figure 1-7 Dielectric failure during chip attachment [20] 10

Figure 2-1 Schematic of a Flip Chip Package 12

Figure 2-2 Signal delay vs technology node [56] 13

Figure 2-3 Flip Chip Process Flow 14

Figure 2-4 Scale difference between various components 16

Figure 2-5 Effective Block Approach (compact modeling) 17

Figure 2-6 Sub-modeling concept in a pulley hub 19

Figure 2-7 Lumping to determine material properties 20

Figure 2-8 Anand’s viscoplastic material model constants for Sn3.5Ag 21

Figure 2-9 Global compact model - package level 24

Figure 2-10 CPI analysis on BEOL stack using detailed 3D modeling of bump cell 25

Figure 2-11 Differences in peeling stress for linear vs non-linear properties for effective block 26

Figure 2-12 Validation of effective properties using warpage 27

Figure 2-13 Validation of effective properties using detailed stress comparison 28

Figure 2-14 Validation of effective properties using detailed strain comparison..... 29

Figure 3-1 Typical FCBGA package 35

Figure 3-2 Full model - 31x31 array..... 35

Figure 3-3 Octant model 36

Figure 3-4 Octant model meshed..... 36

Figure 3-5 Detailed section of the solder ball 37

Figure 3-6 Effective block approach..... 38

Figure 3-7 Cyclic stress-strain hysteresis loop 44

Figure 3-8 Schubert's energy based model for SAC and SnPb solder 45

Figure 3-9 Syed's energy based model for CSPs and BGA 46

Figure 3-10 Reflow loading..... 47

Figure 3-11 Corner joint strain – Array model 49

Figure 3-12 Corner joint strain – Linear "effective block" compact model 50

Figure 3-13 Corner joint von mises stress – array model 51

Figure 3-14 Corner joint von mises stress – linear "effective block" compact model 52

Figure 3-15 Case 1 – Compact model warpage (SAC 305)..... 55

Figure 3-16 Case 1 – Array model warpage (SAC 305)..... 55

Figure 3-17 Case 1 – corner joint von-Mises stress (Array Model SAC 305) 56

Figure 3-18 Case 1 - Corner joint von Mises stress (compact model - SAC 305) 57

Figure 3-19 Case 1 – Equivalent total strain (Array model SAC 305)..... 58

Figure 3-20 Case 1 – Equivalent total strain (compact model SAC 305)..... 59

Figure 3-21 Case 3 - Corner joint von-Mises stress (Array model SAC 305)..... 61

Figure 3-22 Case 3- corner joint von-Mises stress (Compact model SAC 305) 62

Figure 3-23 Case 3 – Corner joint equivalent total strain (Array model SAC 305) 63

Figure 3-24 Case 3 – Corner joint strain (compact model SAC 305)..... 64

Figure 3-25 Warpage variation with solder mask opening..... 66

Figure 3-26 Corner joint stress variation with mask opening..... 67

Figure 3-27 Corner joint strain variation with mask opening..... 67

Figure 3-28 Warpage comparison between “array” and “compact” model..... 69

Figure 3-29 Stress comparison between “array” and “compact” model..... 69

Figure 3-30 Strain comparison between “array and “compact” model 70

Figure 3-31 Correlation model for s_o..... 71

Figure 3-32 Correlation model for S^..... 71

Figure 3-33 Array model - von-Mises stress (mold E = 5GPa)..... 73

Figure 3-34 Compact model – von-Mises stress (mold E = 5GPa)..... 73

Figure 3-35 Array model – Equivalent total strain (mold E = 5GPa)..... 74

Figure 3-36 Compact model – Equivalent total strain (mold E = 5GPa)..... 74

Figure 3-37 Array model – von mises stress (mold E = 25GPa)..... 74

Figure 3-38 Compact model – von Mises stress (mold E = 25GPa) 74

Figure 3-39 Array model – Equivalent total strain (mold E =25GPa)..... 75

Figure 3-40 Compact model – Equivalent total strain (mold E – 25GPa)..... 75

Figure 3-41 Effect of overmold stiffness on warpage 76

Figure 3-42 Effect of overmold stiffness on corner joint strain..... 77

Figure 3-43 Effect of overmold stiffness on corner joint stress..... 77

Figure 3-44 Accelerated thermal cycling test profile 79
Figure 3-45 Array model Warpage 80
Figure 3-46 Compact model warpage 80
Figure 3-47 Array model – corner joint von Mises stress 81
Figure 3-48 Compact model – corner joint von Mises stress 82
Figure 3-49 Array model – corner joint equivalent strain 83
Figure 3-50 Compact model – corner joint equivalent strain 84
Figure 3-51 Array model – von Mises stress (Sn63Pb37 alloy) 87
Figure 3-52 Compact model – von Mises stress (Sn63Pb37) 88
Figure 3-53 Array model – Equivalent total strain (Sn63Pb37) 89
Figure 3-54 Compact model – Equivalent total strain (Sn63Pb37) 90
Figure 4-1 2D vs 3D 92
Figure 4-2 3D innovation roadmap - GLOBALFOUNDRIES 95
Figure 4-3 Stress affects transistor performance [39] 96
Figure 4-4 Localized stress affects the BEoL stack [39] 97
Figure 4-5 TSV Cu protrusion [39] 97
Figure 4-6 Mobility variation with temperature/doping concentration 99
Figure 4-7 Package footprint [57] 101
Figure 4-8 Quarter symmetry model – symmetry faces 102
Figure 4-9 Quarter symmetry model – internal view 102
Figure 4-10 Package Cross-section 103
Figure 4-11 Corner region detailed view 103

Figure 4-12 Sub-model (corner region) 104

Figure 4-13 Array Model cross-section 106

Figure 4-14 Compact model cross-section schematic 107

Figure 4-15 Warpage Comparison (top view) 108

Figure 4-16 Warpage comparison (front view) 109

Figure 4-17 Warpage – compact model (front view) 109

Figure 4-18 Global model TSV von mises stress distribution 110

Figure 4-19 Principal stress – SiO_2 region (sub-model) 110

Figure 4-20 Von-mises stress in TSV (Cu region – sub model) 111

Figure 4-21 Solder μ-bumps von mises stress (Array model – MPa) 112

Figure 4-22 Solder μ-bumps von mises stress (compact model – MPa) 112

Figure 4-23 Solder μ-bumps total strain 113

Figure 4-24 Temperature distribution (no TSV current) 117

Figure 4-25 Temperature distribution 117

Figure 4-26 Junction temperature variation with TSV current 118

Figure 4-27 Performance degradation with TSV joule heating 120

Figure 4-28 Variation in performance hit with chip power for 40mA case (normalized to the baseline case) 120

Figure 4-29 Effect of TSV thermal conductivity on channel noise 121

Figure 4-30 Effect of TSV thermal conductivity on e^- mobility 121

Figure 4-31 Effect of TSV thermal conductivity on transconductance 122

LIST OF TABLES

Table 2-1 Modified material parameters (ANAND) for C4 μ-bumps 23

Table 2-2 Modified material parameters (ANAND) for C4 μ-bumps 23

Table 2-3 Validation summary for the compact model 30

Table 3-1 Material Properties [58, 59, 60] 39

Table 3-2 Material properties for solder [58, 59, 60] 39

Table 3-3 Package Dimensions (full model) 41

Table 3-4 Design of Experiments – Different solder footprints 42

Table 3-5 Mold Compounds with different Elastic Modulus (case 1 from Table 3-4) 43

Table 3-6 Different Solder Alloys – SAC 305 and SnPb 43

Table 3-7 “Linear effective block” vs “Array” configuration 53

Table 3-8 Summary of results – Different solder footprint 65

Table 3-9 Effect of solder mask opening on the mechanical response 66

Table 3-10 Modified viscoplastic model constants for the effective block 68

Table 3-11 Results Summary – Effect of overmold stiffness 76

Table 3-13 Accelerated Thermal Cycling Results 85

Table 3-14 Accelerated Thermal Cycling Results – Damage Correlation 85

Table 3-15 Results summary for Sn63Pb37 solder alloy 90

Table 4-1 Material Properties 105

Table 4-2 Package Dimensions 105

Table 4-3 Summary of Results 114

Table 4-4 Summary of Results – strain validation 114

Chapter 1

INTRODUCTION AND OBJECTIVE

1.1 Role of Packaging in Micro-Electronics

Packaging plays a vital role in any electronic device from the performance and cost standpoint. Packaging provides a medium for the following [7] (see Figure 1-1). It is the whole package that is shipped and not just the silicon; packaging significantly contributes to the total cost - equal to or greater than that of the silicon.

- Interconnect scaling – nm (chip) to cm (PCB)
- Heat dissipation from the device
- High speed signaling
- Mechanically housing the device – protection from environment

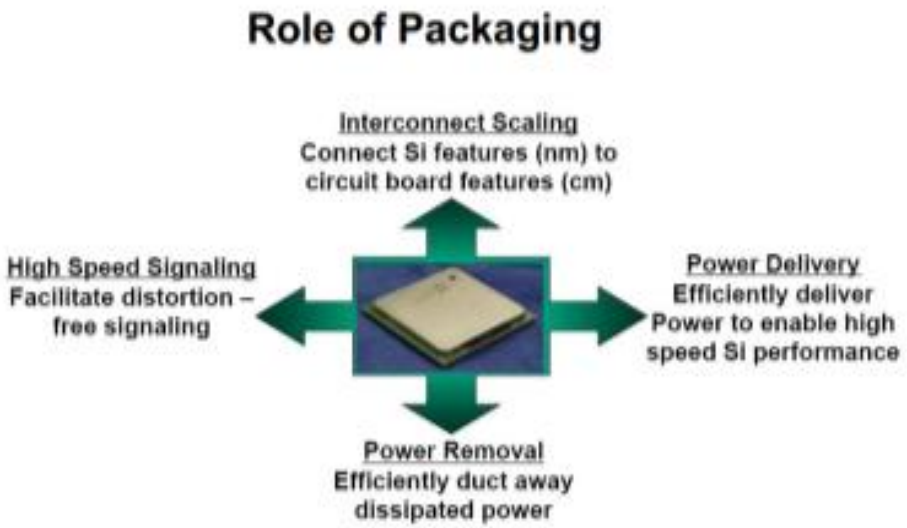


Figure 1-1 Role of Packaging

1.2 Motivation/Literature Review

Since the invention of a transistor in 1947 and development of integrated circuits (ICs) in 1959, there has been an innovative progress in the field of micro-electronics packaging. The key packaging breakthroughs have been coming almost every decade: PTH, leading to surface mount; peripheral array leading to area array packaging and the most significant breakthrough being migrating from planar to 3D packaging. In the last decade or so, the convergence of the consumer electronic products such as cell phones and digital cameras has led to the continued miniaturization of the transistors/ICs, resulting in unparalleled growth of the electronics industry. Since 1970, the number of components per chip has doubled every two years. This historical trend is known as "Moore's Law" [8]. Moore's law has been driving the micro-electronics industry to achieve high performance with small form-factors at a reasonable cost. This has resulted in a reduction of the relative manufacturing cost per function enabling the production of more complex circuits on a single semiconductor substrate [8]. There has been a drastic reduction in cost per bit. In 1954, five years before the IC was invented, the average selling price of a transistor was $5.52. Fifty years later, in 2004, this had dropped to a billionth of a dollar. A year later in 2005 the cost per bit of dynamic random access memory (DRAM) was one nano dollar (one billionth of a dollar) [9] . *As the number of components (i.e. transistors, bits) per chip increases, the total chip size has to be contained within practical and affordable limits. This can be achieved by a continuous downscaling of the critical dimensions in the IC, which can be expressed in terms of Moore's Law as a scaling by a factor of 0.7 (½√2) every 2 years, where "critical*

dimension" is understood as "half pitch", as defined in the International Technology Roadmap for Semiconductors" [10] .

As a consequence of this trend, the miniaturization of circuits by scaling down the transistor has been the principal driver for the semiconductor technology roadmap for last forty years. Moore's law has been the engine of a virtuous cycle (see Figure 1-2).

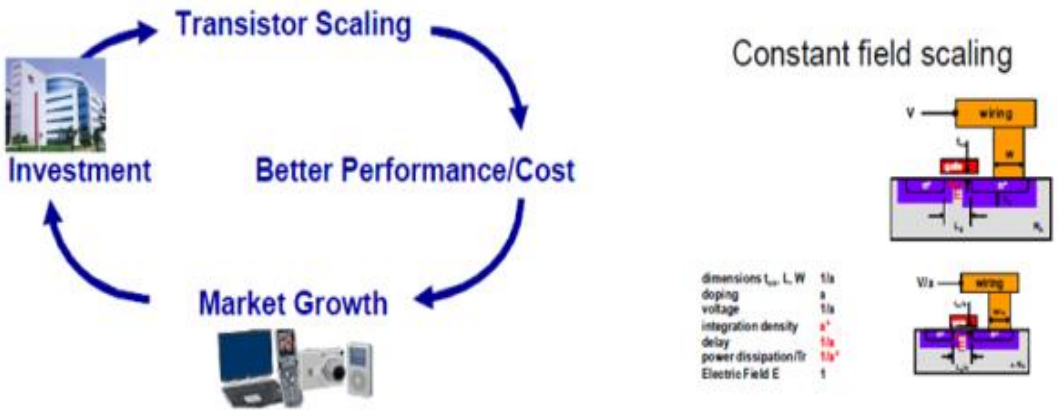


Figure 1-2 Virtuous Cycle and constant component scaling [9]

Migration to a higher technology node provides significant power, performance, and cost improvements; however as shown below, (see Figure 1-3) it is clear that further dimensional scaling is no longer associated with lower average cost per transistor. The chart below, published by IBS about a year ago, shows the diminishing benefit of cost reduction from dimensional scaling. In fact, the chart indicates that the 20-nm node might end up costing higher than the previous node [1].

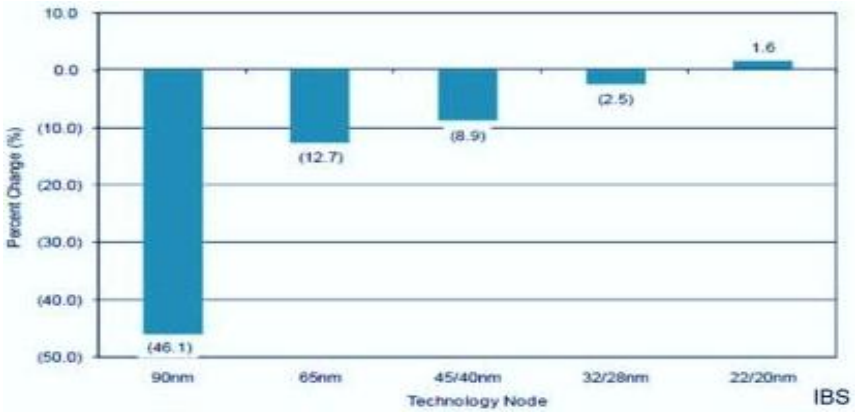


Figure 1-3 Percent change in cost/gate [1]

The cost reduction via dimensional scaling is a result of doubling the number of transistors per wafer. But if the wafer cost of the new technology node increases significantly (see Figure 1-4), then it neutralizes the original cost reduction. The *Nvidia* chart below (see Figure 1-4) shows the wafer cost of recent technology nodes with time. From 80nm-40nm, the incremental wafer cost increase was small, and rapid depreciation of those costs resulted in almost constant average wafer price. However; in recent nodes (28nm, 20nm, 14nm...), the incremental wafer cost has been huge [1]. The root cause of this significant increase in the wafer costs is the increase in the equipment cost required for processing the next technology node - increase in costs of capital, process R&D, and design. To achieve further performance enhancement at a reasonable cost, advanced packaging is being leveraged. Heterogeneous integration and (either SoCs or 3D ICs) is one of the ways to be cost effective. Reduced die size will enable better yield, thereby adding towards the cost effectiveness [1]. Advanced packaging technologies such as FC SOCs, SiPs, and 3-D TSV provide significant power and performance enhancement without having the need to migrate to the higher technology node – basically node

migration can be delayed by few years without compromising performance. FC SoCs are very common in the mobile electronics market because of their low power consumption. A typical application is in the area of embedded systems [11]. When it is not feasible to construct a SoC for a particular application, an alternative is a SiP comprising a number of chips in a single package.

In large volumes, SoC is believed to be more cost-effective than SiP since it increases the yield of the fabrication and because its packaging is simpler [12].

Apple A5X is a 32-bit system-on-a-chip (SoC) designed by Apple, introduced at the launch of the third generation iPad in March, 2012; it is based on the 45-nm technology node. Apple claims it has twice the graphics performance of the A5 [13]. Apple A6X is another large die 32-bit SoC designed by Apple on the 32-nm platform. It is a high-performance variant of the Apple A6. Apple claims the A6X has twice the CPU performance and up to twice the graphics performance of its predecessor, the Apple A5X [14].

Figure 1-4 Wafer cost increment with technology node [1]

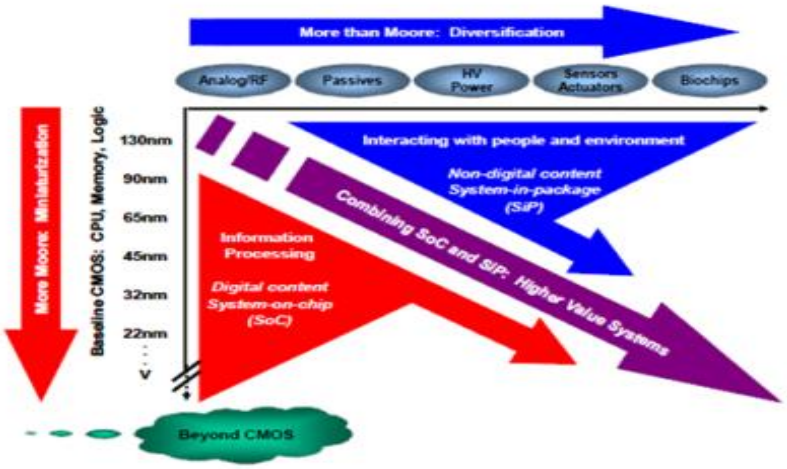


Figure 1-5 “More Moore" and “More-Than-Moore” (MtM)” [9]

As planar device miniaturization continues to its ultimate limits, the complexity of circuit interconnections for 2-D devices becomes a limitation for performance and drives up power dissipation [2]. With increase in die size and circuit complexity (see Figure 1-6), cost and time to market for SoCs drastically increases [15]. Chip-stacking (3-D) is

emerging as a powerful technology that satiates such Integrated Circuit (IC) package requirements. Access to the third dimension has significantly simplified chip-level communications and transfer of information among the processing elements and has provided rapid access to memory and configurable logic. 3-D technology would enable extremely dense solid-state memory to be arrayed within a few microns of the processing elements, which reduces access times.

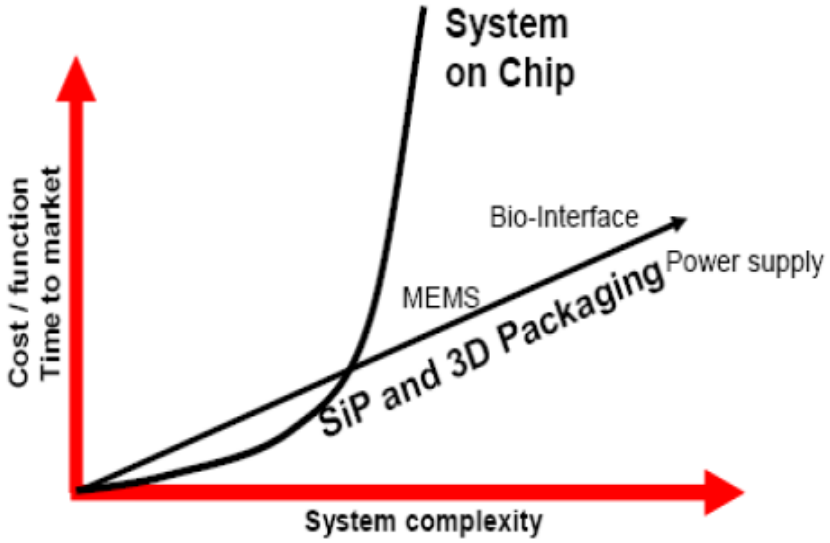


Figure 1-6 SiP/3D and SoC comparison for cost/function/time with system complexity

The 3-D arrangement also provides opportunities for new circuit architectures based on the geometrical ability to have greater numbers of interconnections among multi-layer active circuits. A 3-D FPGA would overcome the interconnect limitations, resulting in greater silicon efficiency per function (number of used gates/total number of gates), faster signal/data throughput, and faster switching of the gate-level configuration. True 3-D integrated circuits can operate at higher clock rates and can consume less power

over their 2-D implementations, as the 3-D arrangement minimizes the length of circuit interconnects [2]. Stacking 24 devices using special etching technology that connects the layers electronically by punching holes from the highest layer to the bottom layer was developed by *Samsung* in 2013 resulting in a higher density 3-D V NAND flash memory [16]. Samsung's 3D V-NAND is able to provide over twice the scaling of 20-nm class planar NAND flash [16]. High-density-interconnects (HDI), necessitates area array packaging with much higher number of interconnects as well as reduced footprints leading to 3D TSV technology. *Micron's* hybrid-memory-cube is one such development [17]. HMC is a small, high-speed logic layer that sits below vertical stacks of DRAM die that are connected using TSV interconnects.

3-D stacking of the processor and memory components in high computing applications reduces the communication delay in a multi-core system owing to reduced system size and shorter interconnects [18]. The shorter interconnection length between the processing and memory components in a multi-core system lowers the overall system access latencies and boosts the system performance. "*Comparing with CSP (chip scale package), which has a silicon efficiency of about 80%; 3-D stacked packages are around 300% without increasing the thickness or the footprint of the package"* [19]. It is clear that the aforementioned advanced packaging technologies are advantageous from cost and performance standpoint, and are therefore shaping up as an integral part of the mainstream consumer electronics, especially portable electronics. However, the development and sustenance of such 'big things" requires immense numerical and

experimental evaluations in order to have a design-for-reliability *(DFR)*. Reliability needs to be evaluated and failures need to be identified and mitigated for smooth and effective transition to such new technologies. There is a need to develop sophisticated and compact numerical methodologies for efficient mechanical analysis.

In planar packages, traditionally, the interconnect bumps were vulnerable against the package CTE mismatch and the chip-package-interaction *(CPI)* risk involved failures in the bumps and the bump to substrate/chip interface. Beyond 65 nm technology node, the use of *low-k* and *ULK dielectric* materials in the wafer Back-End-of-Line (BEoL) has resulted in a failure mode of the dielectric cracking (see Figure 1-7) , commonly known as "white bump issue" during flip chip attachment to the substrate [20]. Analysis of "white bump" failure requires full field 3D modeling of the structure (huge scale difference ~ 5 orders of magnitude), thereby causing a bottleneck in the development phase. The huge I/O count further complicates the FE modeling, making it highly computationally expensive and in some cases impossible. So, there is a need to develop a scalable compact modeling methodology for quick and accurate thermo-mechanical predictions in such systems.

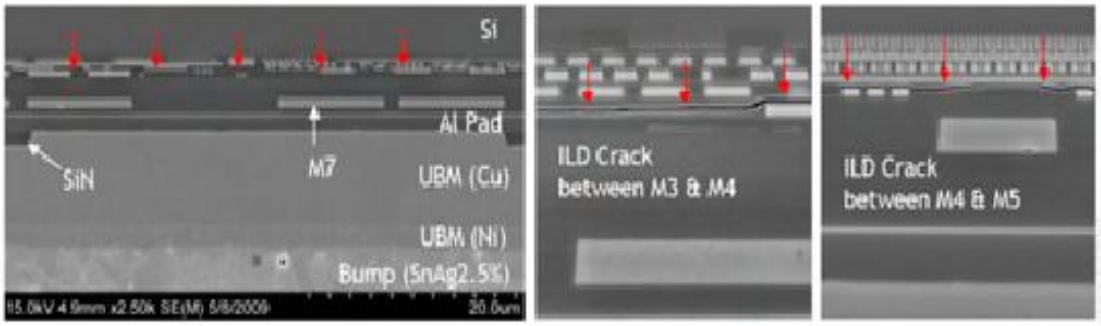


Figure 1-7 Dielectric failure during chip attachment [20]

In this work, a novel compact modeling methodology has been demonstrated to analyze the multi-level interconnects damage during process for planar and 3D TSV packages. Scalability of the proposed modeling methodology across different interconnect levels and packaging technologies has been established through this work. It is shown that the proposed methodology is very useful for the multi-level *(CPI and 2nd level interconnect)* thermo-mechanical analysis as it significantly reduces the computational time without compromising the prediction accuracy. Feasibility study is carried out in 2 steps – 1) *CPI* analysis for *BEoL/fBEoL* damage assessment during the flip chip attachment to the substrate; 2) 2nd level ball-grid-array (BGA) damage assessment during package attachment to the printed-circuit-board (PCB). To validate the proposed methodology, several damage correlation parameters such as maximum warpage and solder bump/joint stress-strain distribution have been compared between the "compact model" and "full array" detailed model. Both qualitative and quantitative comparisons have been shown in order to have immense confidence to leverage the novel methodology in real time systems. Furthermore, compact thermal and mechanical

modeling of a 3-D TSV flip chip package is performed to study the effect of joule heating on the device performance measured by trans-conductance, electron mobility (e^- mobility), and channel thermal noise. Intrinsically, the 3-D TSV modeling is complicated and very time consuming because of large number of miniature I/Os (TSVs and inter-die μ-bumps/pillars), thereby warranting a compact modeling approach. A 2-die 3-D TSV package is leveraged and TSV/SiO_2 interaction during package attachment to the substrate is studied and compared between the "full array" (with TSVs and μ-bumps) and the "compact model" to establish the feasibility of the compact modeling approach in 3-D packages. Preliminary results indicate that joule heating has a significant effect on the thermal response of the 3-D IC and subsequently proves to be quite detrimental to the over-all performance.

Chapter 2

PLANAR PACKAGES – COMPACT MODELING METHODOLOGY DEVELOPMENT TO STUDY CHIP-PACKAGE INTERACTION IN FLIP CHIP PACKAGES

2.1 Chip-Package-Interaction (CPI) Prediction – Why is it a concern?

Flip chip is a proven high-density-interconnect (HDI) with an array of solder bumps/Cu pillars as "interconnections" between the silicon and the substrate (see Figure 2-1). Chip Package Interaction (CPI) is a widely recognized reliability and quality challenge for flip-chip packages due to the low-k and ultra-low-K materials used within the silicon BEOL. CMOS technologies beyond 65nm require use of low-k/ultralow-k porous materials ($k<2.5$) as inter-layer dielectric (ILD) to reduce the resistance-capacitance (RC) delay (see Figure 2-2). However, mechanical properties of these materials deteriorate due to increase in porosity with lower K, thereby raising concerns about the integration and reliability of Cu/low-k interconnects within the silicon BEoL [21].

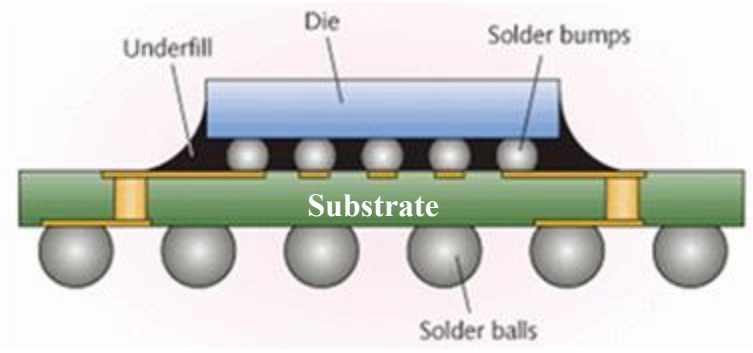


Figure 2-1 Schematic of a Flip Chip Package

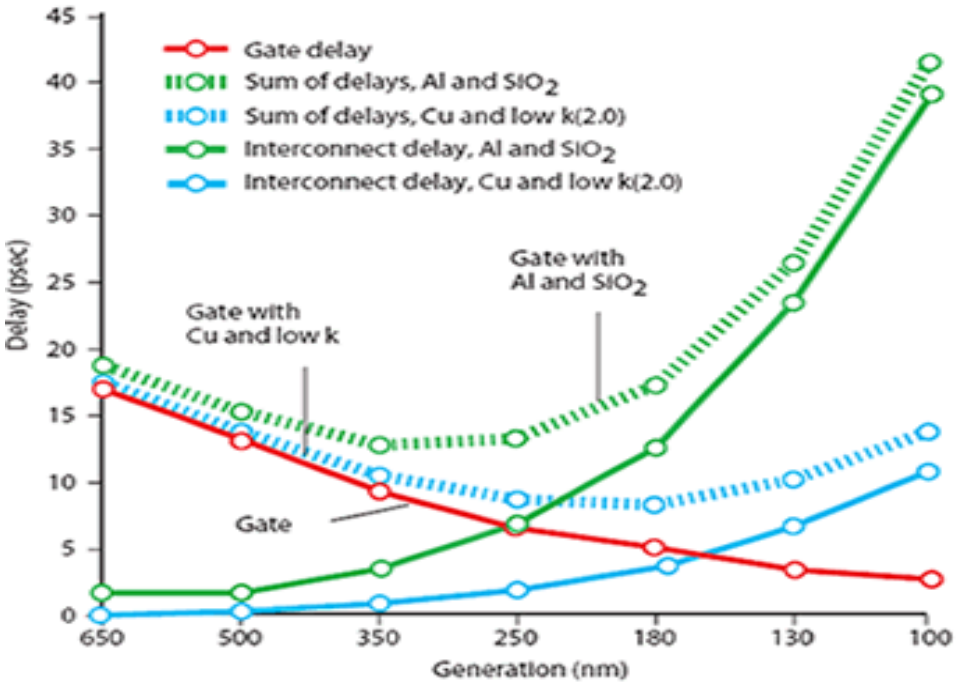


Figure 2-2 Signal delay vs technology node [56]

The sequential IC packaging and assembly processes also pose significant risk to the *BEoL/fBEoL* due to thermo-mechanical stress buildup from CTE mismatch between the silicon and the substrate during the mass reflow solder profile temperature from 260°C to RT (room temperature) [22]. Furthermore, for advanced flip-chip package models, it is challenging to have the entire C4 bump array within the model due to the extremely large pin count (I/O) fitted at a fine pitch within a large die. As an example, for CPU chips, it is not uncommon to have chip contain over 5000 C4 bumps. Accounting for such large bump count with FEA models, makes the analyses not just computationally expensive but also often impossible. The proposed study will try to address these challenges by demonstrating a "compact" 3D modeling approach for *BEoL/fBEoL* region.

Flip chip assembly process shown below (see Figure 2-3). Since there is no underfill (during the flip chip attachment to substrate) to bolster the *BEoL/fBEoL* region, the mechanical risk is even higher and could lead to fracture or delamination in the dielectric layers and/or solder bump during attachment (see Figure 2-3). To mitigate these risks, finite element (FE) modeling of CPI is practiced during product development to identify the design weaknesses early on and propose alternatives.

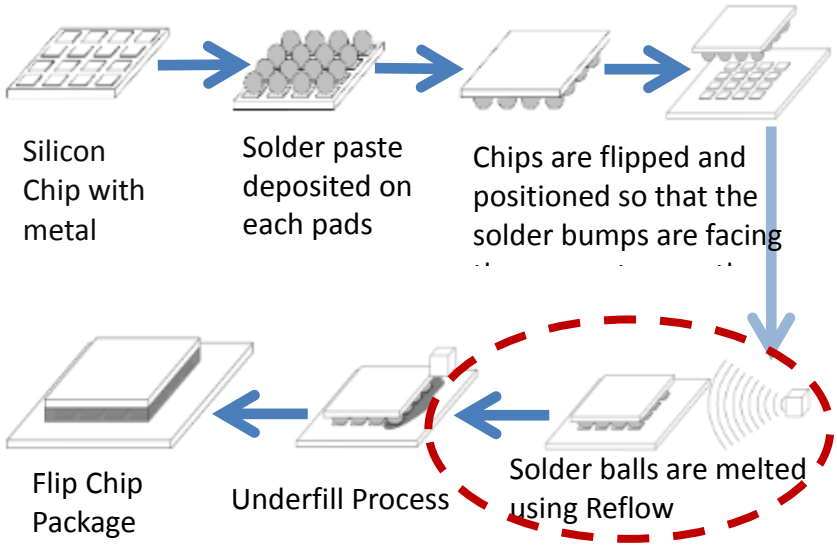


Figure 2-3 Flip Chip Process Flow

2.2 Literature Review

Many studies on solder bump reliability in flip chip technologies have been extensively reported in literature. Liu *et al*. [23] and Chiang *et al.* [24] found that solder joint reliability is highly dependent of solder joint geometry, such as their standoff height, lower/upper contact angles of solder joints, and solder joint profile etc. Popelar [25] studied the reliability of a flip chip package based on FE parametric analysis. He concluded that the flip chip packages without underfill show worse reliability and the design parameters such as die size, solder bump height, and pad diameter would substantially influence the reliability of the packages. Popelar [25] showed that the flip chip packages with underfill show better reliability, and these design parameters (die size, solder bump height, and pad diameter) show smaller influence to their reliability.

2.3 CPI Assessment on BEoL/fBEoL stack of large I/O chip

For advanced flip-chip packages involving large die, high pin-counts (I/O) and finer pitch, the scale differences between package and chip level features could be up to 5 orders of magnitude (see Figure 2-4), which makes the Chip-Package Interaction (CPI) analysis computationally challenging or often impossible. Literature research suggests that former studies have adopted either a "2D approach" or a "reduced 3D strip approach" [26, 27] to address CPI risks for such form-factors. Neither of these techniques accurately represents the real 3D stress state existing in IC package components due to over-simplification of the geometry and/or boundary conditions. The proposed study

addresses these challenges by developing a compact 3D multi-level FE modeling methodology to address the CPI for 28nm technology node.

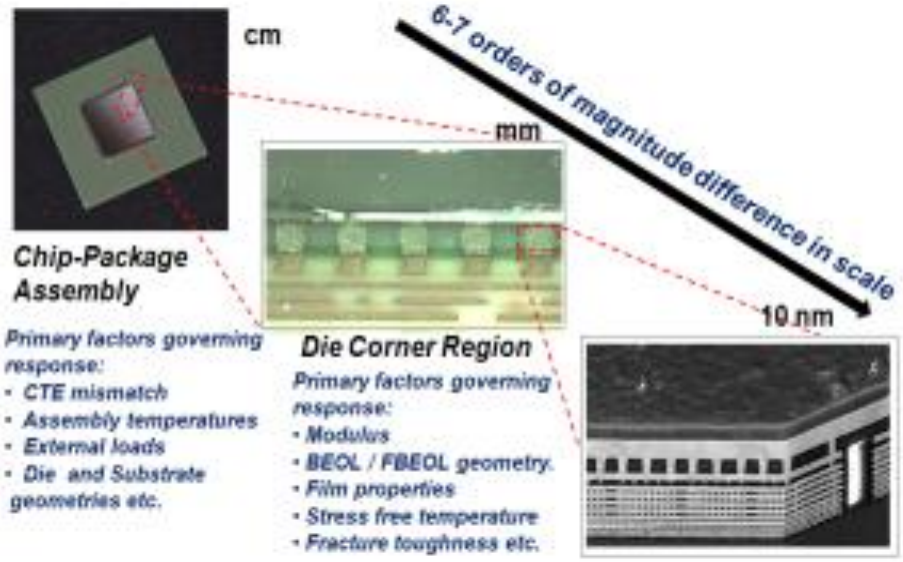


Figure 2-4 Scale difference between various components

2.4 Compact 3D FEA Modeling

The bumps at the geometric corner region of the die have higher stresses due to the CTE mismatch forces being proportional to the distance from neutral point DNP. Consequently, the BEoL layers underneath the corner region bumps are seen to be more susceptible to cracking/delamination under thermo mechanical loads. Hence, as shown (see Figure 2-5), this methodology adopts a "balanced" approach by having the die corner region modeled in detail with actual bump geometries, pitch, pad openings, material sets etc. To address the challenges with large I/O, a non-linear material model is developed based on the Anand's viscoplastic model [28]. This model represents the entire bump

array away from the corner region of the die, whereas the corner region has a detailed bump map of the actual package. Since the bumps at the geometric corner region of the die have higher stresses due to their distance from neutral point DNP, the BEoL layers underneath them are more susceptible to cracking/delamination [6]. The developed methodology has been rigorously validated against the full-array configuration for critical damage parameters and offers up to 10x reduction in the compute time while maintaining accuracy similar to that of the full-array configuration.

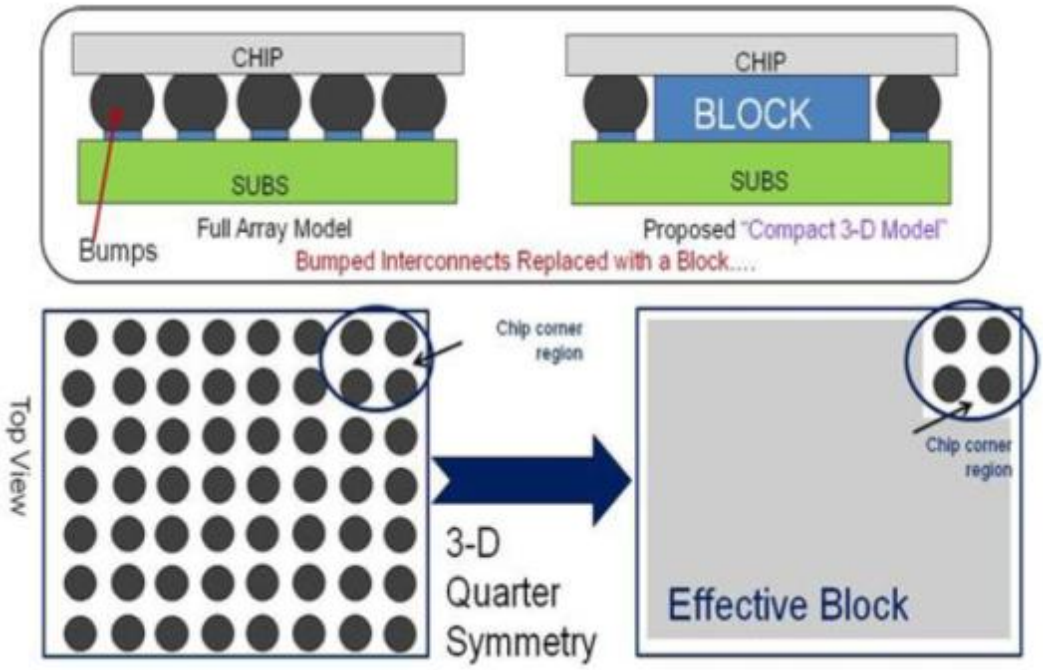


Figure 2-5 Effective Block Approach (compact modeling)

Modeling has been performed in 2 steps; 1) global model – full field 3D global model; 2) sub-model – fine meshed far corner unit cell (shown in Figure 2-10). In FE analysis it may occur that mesh is too coarse to provide satisfactory results in the area of interest where stress in higher, also there could be some very small features that cannot be modeled in the global model. Sub-modeling technique is usually leveraged for such situations for efficient and pragmatic analysis. Sub modeling is sometimes known as global-local analysis or cut boundary displacement method. Cut boundary is the boundary of the sub model where it has been cut through global model. Displacement calculated on the boundary of the cut from global model is applied as a boundary condition for the sub model at cut boundary planes. Figure 2-6 explains how area of interest (high stress) from global model of pulley hub and spokes is differentiated in sub model.

St. Venant's principle supports the sub-modeling technique. It states that if actual force distribution is replaced by statically equivalent system, the distribution of stress and strain is altered only near the region of load application. This explains that stress concentration effects are localized around the concentration, so if the boundary of the sub model is far enough away from the area of interest, reasonably accurate results can be calculated in the sub model.

Apart from just the accuracy there are other benefits of sub modeling, which are stated below:

- The need for transition region in solid FE models is reduced or eliminated.
- It allows you to experiment on different design and area of interest.

- It helps you in getting adequate mesh refinement.
- Add more features/geometries which are impractical to model in the global model due to size difference.

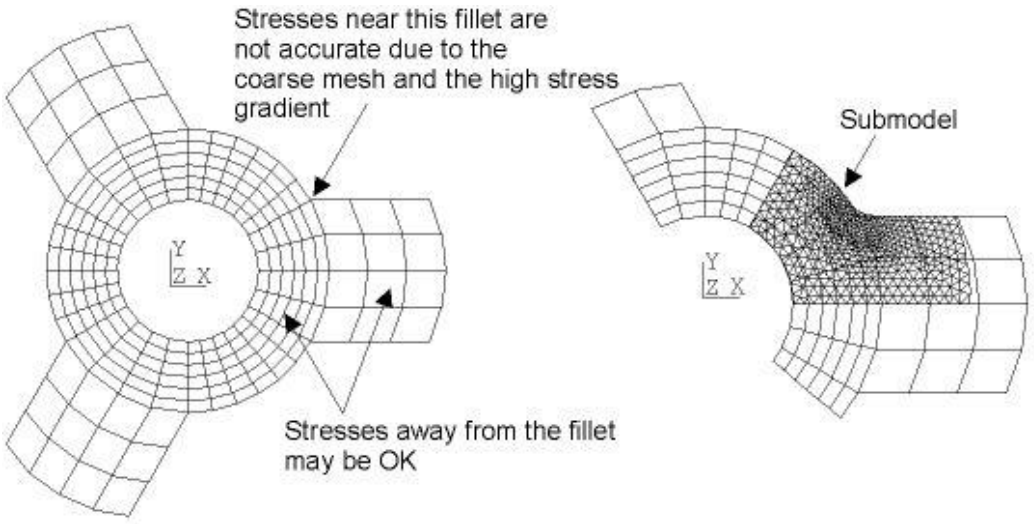


Figure 2-6 Sub-modeling concept in a pulley hub

2.5 Linear property evaluation for compact modeling

The concept of calculating effective properties has been leveraged by many researchers for linear isotropic material models. This is by using fundamental principles of mechanics; it is possible to break down a material cross-section into a series and parallel configuration to arrive at effective properties. Figure 2-7 shows one such example where the chip to substrate interconnects is solder μ-bumps. These interconnects are commonly used for fine pitch, large I/O packages for electrical, mechanical and

process benefit. Series element calculation assumes that the spring elements share the same load for a combined displacement of each material in series. Whereas, the parallel element calculation assumes that under the action of load, the materials displace by the same amount while sharing the total load fraction. For chip package without the underfill, the parallel calculation for the bump cell needs to account for the air gap that exists. Coefficient of Thermal Expansion, CTE and modulus, E value of zero is recommended to be chosen for the parallel calculations. Based on the series and parallel calculations for different materials involved (shown in Figure 2-7) the effective linear properties of the "effective block" have been determined and deployed in the material models. Since the effective block is replacing the μ-bump array (viscoplastic material), a "linear only" representation may prove to be a crude assumption/simplification of the material model. So, to incorporate the inelastic material behavior, non-linear material model for the effective block was determined by tuning the constants such that the compact model mechanical response matches the full array response.

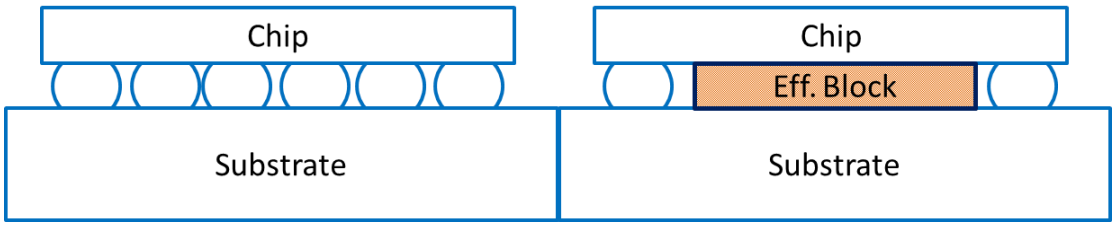


Figure 2-7 Lumping to determine material properties

$$E_{eff} = E_s V_s + E_{air} V_{air} \qquad \text{Equation 2-1}$$

$$\alpha_{eff} = \frac{\alpha_s E_s V_s + \alpha_{air} E_{air} V_{air}}{E_s V_s + E_{air} V_{air}} \qquad \text{Equation 2-2}$$

E_s = Young's Modulus for solder (GPa)

E_{air} = Young's Modulus for air (GPa) = 0

V_s = Volume fraction for Solder

V_{air} = Volume fraction for air

α_s= Coefficient of thermal expansion for solder bumps/balls

α_{air} = Coefficient of thermal expansion for air = 0

α_{eff} = Coefficient of thermal expansion for the effective block

E_{eff} (block) = Young's Modulus for the effective block

MATERIAL PARAMETERS OF ANAND MODEL FOR LEAD FREE SOLDER Sn-3.5Ag

Material parameter	A	Q	$\hat{s}$	h_0	ξ	m	n	a
	sec^{-1}	J/mol	MPa	MPa				
value	177016	85459	52.4	27782	7	0.207	0.0177	1.6

TABLE III
INITIAL VALUES OF INTERNAL VARIABLE s FOR LEAD FREE SOLDER Sn-3.5Ag

Temperature (K)	233	253	313	353	398
s_0 (MPa)	13.9925	9.6351	9.5253	3.0417	2.3165

Figure 2-8 Anand's viscoplastic material model constants for Sn3.5Ag

2.6 Non-linear property evaluation for compact modeling

Another factor that is important to be considered here is the non-linearity of solder. Solder is a viscoplastic material and rate-dependent plasticity (ANAND) model is chosen as a standard in the industry for representing the behavior of solder under thermo mechanical loading. Figure 2-8 shows the typical material parameters for Sn-3.5Ag solder as referenced by Chen et.al. [28] .

There isn't any series/parallel algorithm like used in linear model above for calculating effective properties using non-linear material constants. Additionally, even if rule-of-mixture type calculations were done, such models have been found to not truly represent the thermo mechanical behavior of the bump array which the effective block is supposed to represent. So, to overcome this, the tuning for nonlinear effective properties is done as explained earlier using thermo mechanical response as a criterion.

Table 2-1 below shows the tuned parameters after several iterations to match the thermo- mechanical response of the package with effective properties to that with the full bump array. As can be seen in the tables below, all constants except $\hat{S}$ and s_0 can remain unchanged from table 1. $\hat{S}$ is the "coefficient of deformation resistance", whereas s_0 is the "initial deformation resistance- initial value of $\hat{S}$" of the solder material. By tuning the deformation resistance at various temperatures, we have arrived at the most optimum set of properties that match very well with thermo mechanical response of the package. The next section discusses about how these properties have been used in the global-local modeling framework.

Table 2-1 Modified material parameters (ANAND) for C4 μ-bumps

(Proposed in this study)

Alloy/Configuration	Q/R	A (sec^{-1})	ξ	m	h_o (MPa)	$\hat{S}$ *(MPa)*	n	a
Non-Linear Block – C4	10278	1.7e5	7	.207	27782	*11*	.0177	1.6

Table 2-2 Modified material parameters (ANAND) for C4 μ-bumps

(Proposed in this study)

	Temperature (K)				
ANAND CONSTANT (s_o) MPa	*233*	*253*	*313*	*353*	*398*
Non-Linear Block – C4	*7.5*	*5.7*	*5.6*	*2.9*	*.95*

2.7 Results and Discussions

To address the challenges with feature scaling, a multi-level thermo-mechanical modeling framework has been developed on 28-nm technology to analyze the stress-strains from package-level down to the bump or the *BEoL* level. The global (package level) model and local (die level) model are shown below (see Figure 2-9 and Figure 2-10). Detailed layers, passivation openings and pad geometries are implemented at the die level model. A die and substrate size of roughly $400mm^2$ and $1600mm^2$ respectively, is used for the simulation. Further details about the global-local modeling approach have been discussed elsewhere [21].

Figure 2-10b) shows the mechanical strain experienced by the corner-most bump of the package. This methodology is scalable across various technology nodes and is used to analyze sequential stress build-up in silicon BEoL layers during flip-chip assembly process.

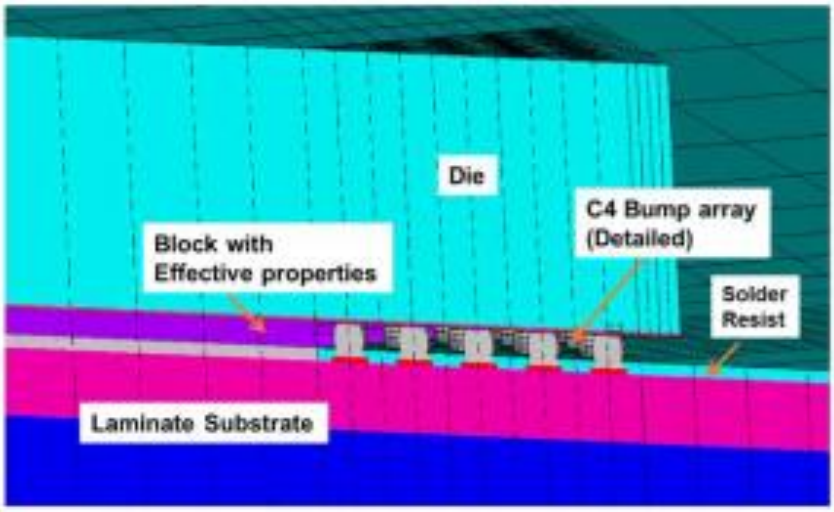


Figure 2-9 Global compact model - package level

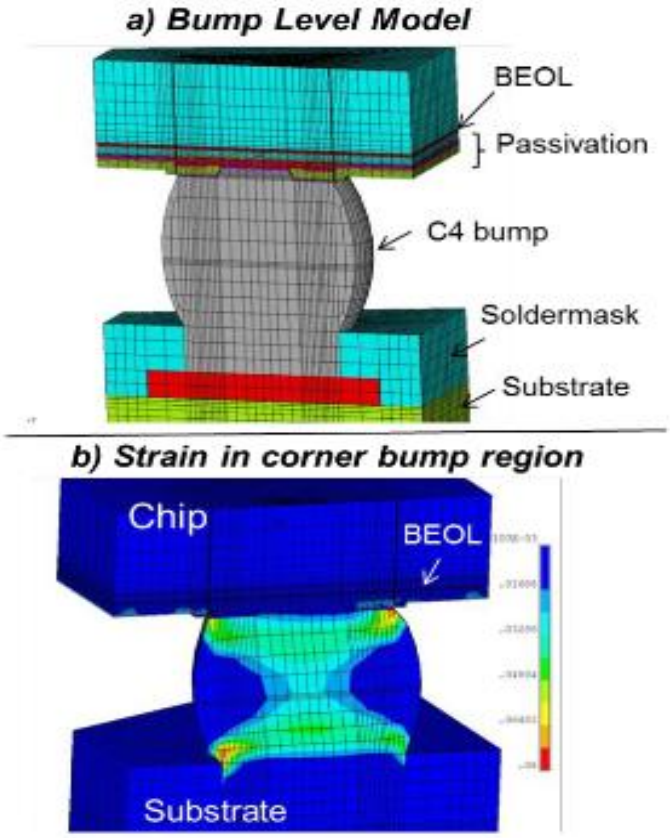


Figure 2-10 CPI analysis on BEOL stack using

detailed 3D modeling of bump cell

As discussed in the previous section, the effective block can have either linear or nonlinear properties. Using the material model shown in (see Table 2-1), the peeling stress obtained in the interconnect region is shown in (see Figure 2-11). As can be seen, the linear block model does not give a very accurate representation of the stress filed because artificial stress concentration is seen in the interface between the block and the bump array. The nonlinear block on the contrary is more representative of actual stress field.

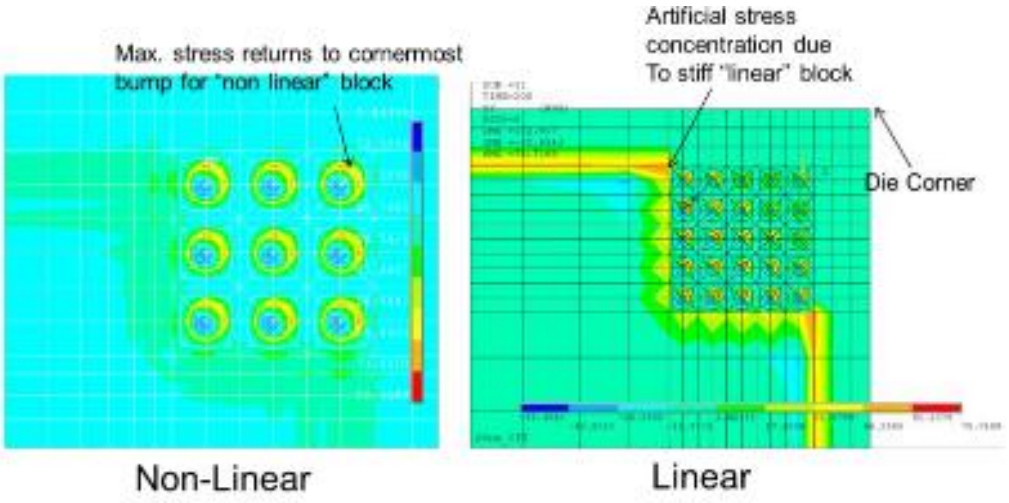


Figure 2-11 Differences in peeling stress for linear vs non-linear properties for effective block

For an apple to apple comparison between the effective block (compact model) and the full array model, an effort is made to match the die and package sizes and plugin the properties evaluated so far. The die size had to be shrunk to 25 mm^2 to reduce the compute time required for simulation. Multiple simulations were done to understand die size effect on the scalability of the effective block technique. Figure below (see Figure 2-12) shows the warpage seen in the package after the chip-attach process. Both models show a very good correlation in terms of out-of-plane displacement or warpage.

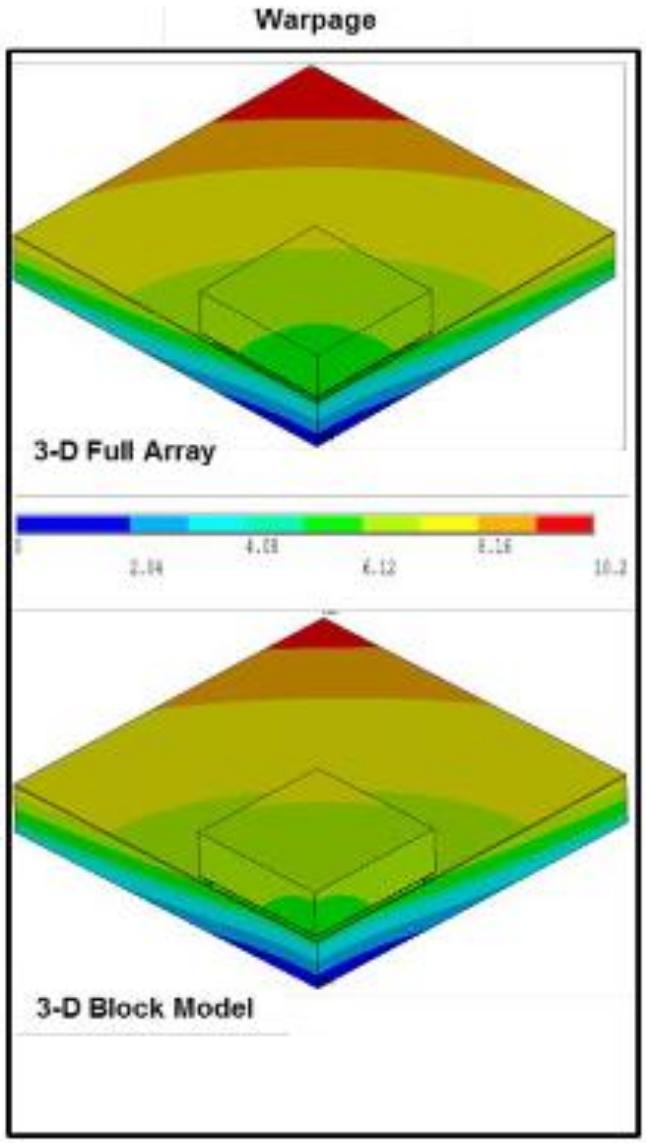


Figure 2-12 Validation of effective properties using warpage

If we cross-section the die corner region then bump array should also show the same stress field across both models. This is verified by the figure (see Figure 2-13 and Figure 2-14) which demonstrates that not only is the warpage correlating well, but also the solder *(fBEoL)* and the underlying BEoL layers that match very well when we

compare both the models. Basically, the "non-linear effective block" has a similar effect on the corner region as that imposed if there were solder bumps.

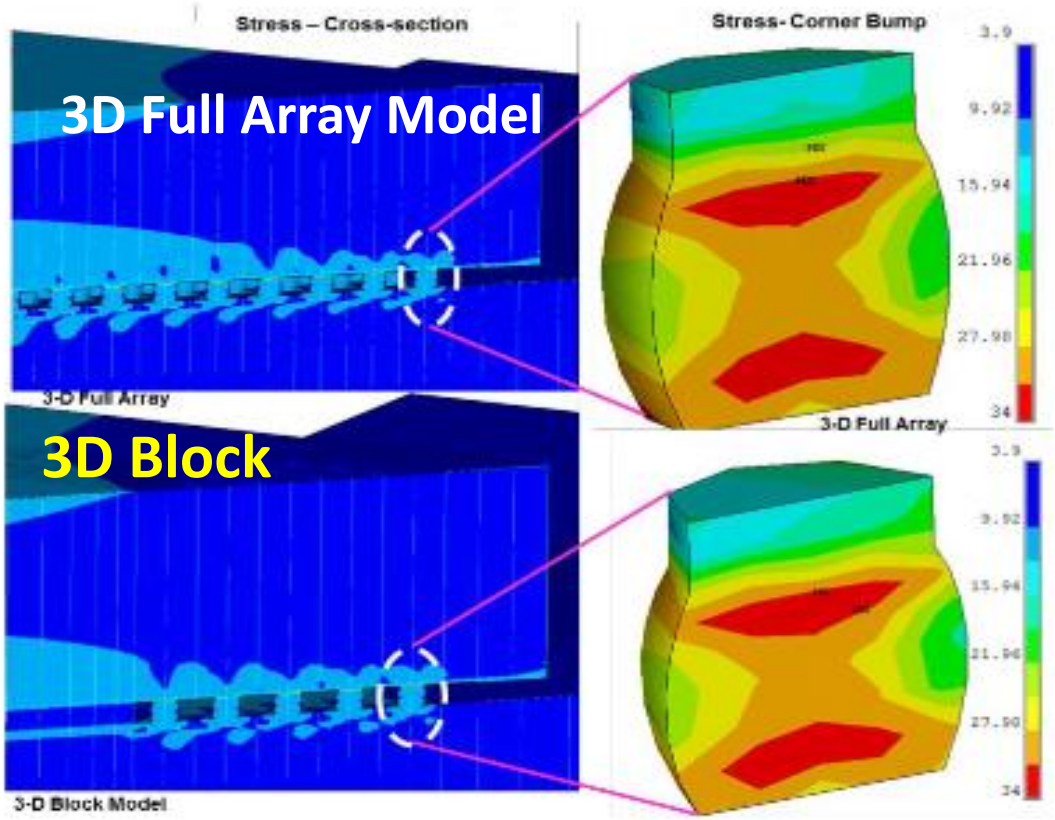


Figure 2-13 Validation of effective properties using detailed stress comparison

The developed methodology has also been validated in the local model where detailed pad geometries and film representations are present. Figure 2-14 shows for one of the inter-layer-dielectric (ILD) layers, the difference in the principal strain values is less than 5%. If we continue looking at other layers such as ULK, a nice comparison between both cases can be done. Table 2-3 summarizes the findings for the C4 µ-bump array effective block model response at different levels *(fBEoL/BEoL)*.

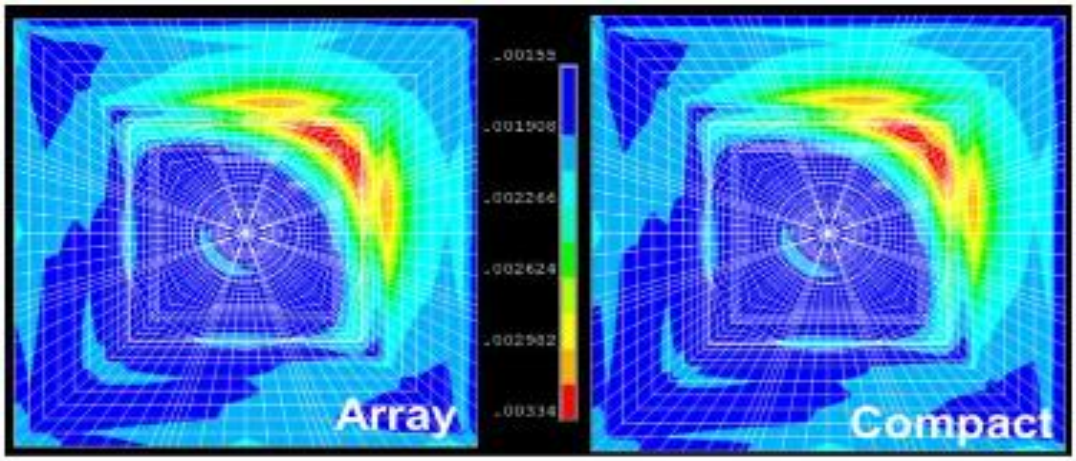


Figure 2-14 Validation of effective properties using detailed strain comparison

The developed methodology can also be used for Cu pillar arrays as mentioned earlier. The material constants summarized in Table 2-1 are chosen such that they are robust across multiple configurations like pitch differences, height differences, die size differences, bump diameter variations etc.

Table 2-3 Validation summary for the compact model

Stress Parameters			
	Array model	Compact model	% Difference
BEoL layer 1 Principal Stress S1 (MPa)	108.3	107.2	1
BEoL layer 2 Principal Stress S1(MPa)	76.3	67.5	11
Solder Bump – Stress (MPa)	33.3	33.4	-
Solder Bump – Strain (%)	5.9	5.9	-
Maximum Warpage (μm)	10.2	9.9	3

A novel 3D compact modeling framework has been developed to perform multi-level CPI risk analysis of a large I/O chip during flip-chip assembly. CTE mismatch forces are highest at regions away from the die center. So a detailed model of the die corner region has been built, while a visco-plastic material model (modified ANAND) has been developed for the rest of the interconnect region (effective block) to preserve the temperature and time rate dependence of the structure. Compact model has been validated against standard stress parameters for a full-scale 3D array model. Qualitative and quantitative correlation work has been presented to have confidence in the accuracy and scalability of this framework. The developed methodology can scale across various technology nodes to mitigate the white-bump risk during chip-package assembly. The

scalability of this technique for laminate substrate to PCB interconnects (2^{nd} level BGA joint) is demonstrated in the next chapter. Lastly, the reduction in numerical compute time is up to ***10x*** or more for a typical configuration. This reduction in compute time comes without a loss in accuracy of model predictions for warpage, stresses and strains in a CPI analysis, leading to immense time savings during reliability development.

2.8 Acknowledgement

This work (chapter 2) was performed during the internship in the "Technology Reliability Development" group at GLOBALFOUNDRIES, Malta, NY in 2012-2013. I sincerely thank Mr. Premachandran C S (*SMTS- QRA*), Chirag Shah (Ex Principal Engineer), and Dr. Natarajan Mahadeva Iyer (Director-QRA) to give me an opportunity and guide me through this project that enabled it to be a worthwhile venture for GLOBALFOUNDRIES and me. This work was published in the *IEEE Electronics Packaging Technology Conference (EPTC)* 2013 held at Singapore.

Chapter 3

PLANAR PACKAGES - SCALABILITY OF THE PROPOSED COMPACT MODELING METHODOLOGY – 2nd LEVEL BGA DAMAGE ASSESSMENT

3.1 Introduction

The continued increasing number of transistors on the chip has led to a requirement of having smaller, finer pitch and increased I/Os at all the levels (0^{th}, 1^{st}, and 2^{nd}). The increased I/Os and finer pitches are useful from the electrical/performance standpoint; however the thermo-mechanical analysis becomes very challenging in such systems due to newer materials, smaller footprints, thinner layers, and so on. Over the last 2 decades or so, finite element (FE) technique has been widely used to analyze and alleviate the thermo-mechanical reliability challenges at the 2^{nd} level interconnection. However; with multi-level miniaturization and fine pitch solder ball arrays, it has become extremely challenging to simulate the thermo-mechanical response of state-of-the-art form-factors within a reasonable computational time. Sub-modeling has been leveraged to keep a tap on the model size and the solution time without making too many simplifications (geometric, constraint, loadings, etc.,) in the region of interest. Sub-modeling technique has significantly eased and enhanced the simulation capabilities but looking at the miniaturization and increased I/Os, even sub-modeling would be relatively time-consuming as it requires preprocessing and solving at the global level before moving on to the sub/local model. Majority of the work in the reliability assessment at

the BGA level uses either a three-dimensional (3D) full array models or 2D models to predict the solder joint life under various temperature loadings [24-28].

3.2 Literature Review

Lin *et al.* [29] performed a non-linear 3D board-level finite-element analysis (FEA) to characterize the second-level interconnect BGA strain distribution difference under power cycling and thermal cycling tests with underfill and substrate core material splits for a large die flip chip package. He demonstrated an acceleration stress methodology for second-level interconnects reliability under power cycling by increasing temperature ramp rate (60°C/min) and raising peak temperature (135°C), while still keeping the same failure phenomenon. Mercado *et al.* [30] performed a finite element (FE) study to analyze the effect of solder pad footprint on the PBGA reliability under thermal cycling. He also studied the effect of substrate thickness on the mean-time-to-failure (MTTF) for the vulnerable solder joint. Lai *et al.* [31] performed a dimension study of the sub model and concluded that if the pitch is taken as width of the sub model, it provides a good estimate of the results in view of the St. Venant's principle. Cheng *et al.* [32] demonstrated a parametric finite element analysis over a number of geometry/material design parameters to investigate the dependence on the fatigue lives of the thermally loaded solder joint in a typical thermally-enhanced (TE) BGA assembly. Extensive work has been performed in the area of BGA reliability prediction in the last 2 decades or so. Gustafsson *et al.* [33] presented a study on the life prediction analysis of BGA packages by considering different linear and non-linear elements. Almost all the

work done so far has used either a three-dimensional (3D) full array configuration or 2D models to predict the solder joint life under various temperature loading conditions.

Shah *et al.* [34] demonstrated a compact modeling technique for back-end of line (BEoL) in flip chip. It was showed through this paper that compact modeling at the 1st level significantly reduces the computational expense with negligible compromise in the prediction accuracy.

In this chapter, a compact modeling methodology similar to that shown in chapter 2 and [34] has been leveraged for mechanical response prediction at the 2nd level BGA interconnects. A typical flip chip ball-grid-array (FC BGA) package has been leveraged for the correlation analysis. Comparative analysis between the "compact model' and the "full-array model" configuration is demonstrated to validate the proposed technique and show that the proposed technique is scalable across multi-level interconnects and different solder alloys. Feasibility of the proposed compact modeling technique is strengthened by qualitative validation with the published literature.

3.3 Compact Modeling and Methodology

A typical FCBGA package is shown in the figure below (see Figure 3-1).

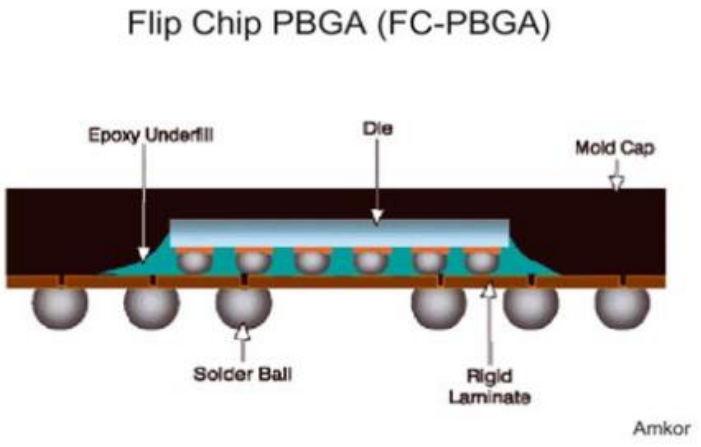


Figure 3-1 Typical FCBGA package

The full model has an array of 31 x 31 solder balls (see Figure 3-2). To save computational time octant symmetry of the full model was formulated as shown in the figure (see Figure 3-3 and Figure 3-4).

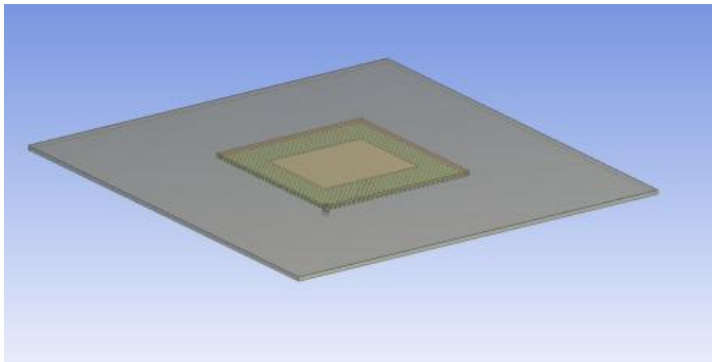


Figure 3-2 Full model - 31x31 array

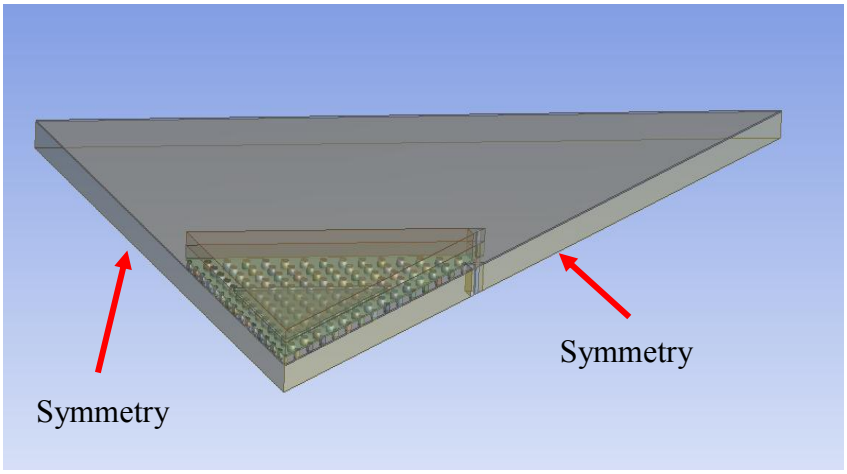


Figure 3-3 Octant model

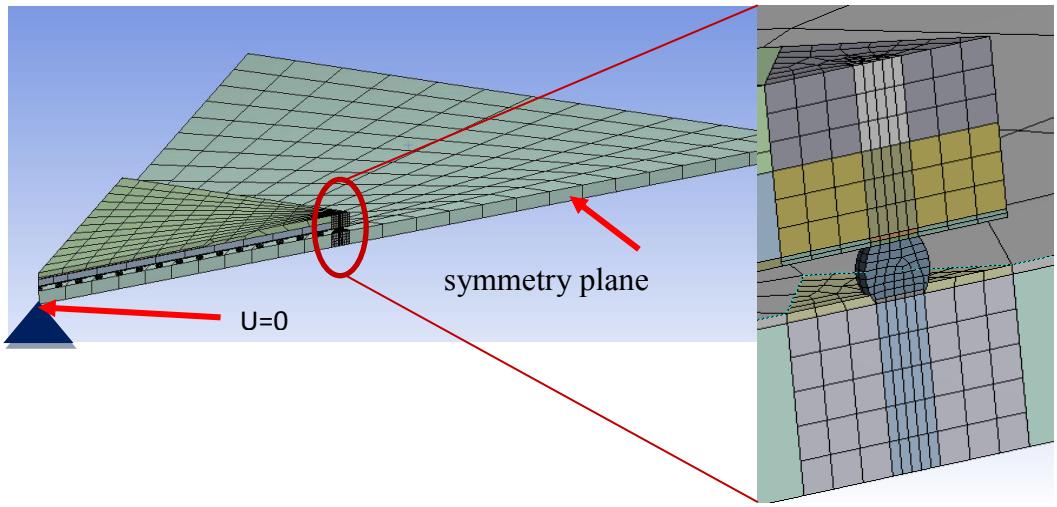


Figure 3-4 Octant model meshed

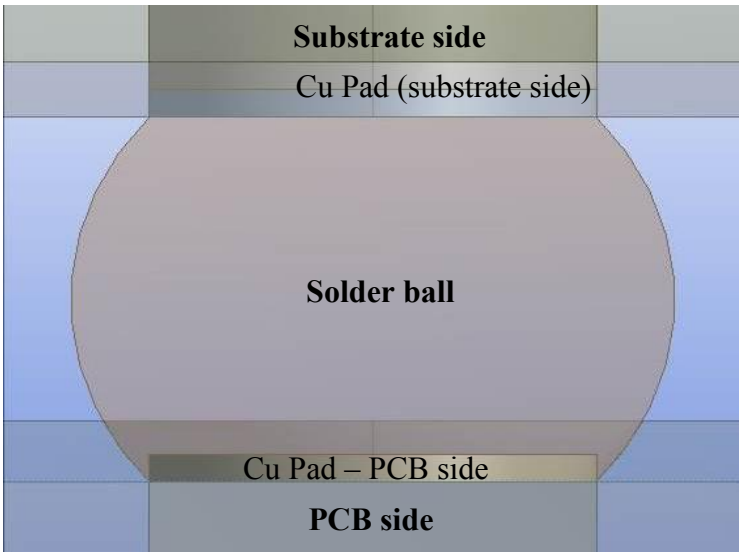


Figure 3-5 Detailed section of the solder ball

Initially the complete model with a full array of solder balls was simulated under package attachment to the PCB (200°C to RT for SAC alloy and 170°C to RT for SnPb) and the values of all the correlation parameters – maximum warpage and the solder ball stress/strain distribution is observed and compared with the compact model. The solution time was also noted. Similar to the effective block used in chapter 2, an effective block representing the inner solder balls has been used (see Figure 3-6). The linear and non-linear material properties for the block were calculated exactly in the same manner as done in chapter 2 for the 1st level μ-bumps. The computational time for the compact model configuration was also noted and compared with the full array model.

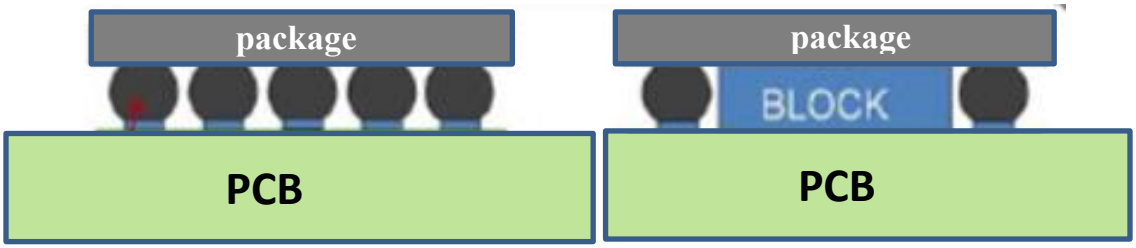


Figure 3-6 Effective block approach

Each layer in the package is assumed to be perfectly bonded to each other. An octant symmetry 3-D model was used in this analysis and one common vertex of the PCB was fixed (all Degrees of Freedom zero) to avoid rigid body motion. Symmetric boundary conditions were applied at the octant symmetry plane. All materials except solder alloy (SAC 305 and Sn63Pb37) are modeled using linear elastic material properties. Both the time and temperature dependent nonlinear properties are used for the solder to capture the inelastic behavior (Anand's constants for SAC 305 and Sn63Pb37).

3.4 Material Properties and Dimensions

All the materials except the solder alloys are considered linear elastic. Substrate and the PCB have orthotropic thermal expansion coefficients. Table 3-1 and Table 3-2 show the material properties for different package components.

Table 3-1 Material Properties [58, 59, 60]

Material	Young's Modulus (GPa)	Poisson's Ratio	Coefficient of Thermal Expansion (ppm/°C)
Die	131	0.28	3
Die Attach	10	0.3	33
Solder mask	4	.4	52
Over Mold	5, 25, 40	0.3	8, 32 (post Tg)
Cu Pad	110	0.34	17
Substrate	24.5	0.22	17 (X), 17 (Y), 52 (Z)
PCB	17.2	0.3	17.6 (X,Y), 64.1 (Z)

Table 3-2 Material properties for solder [58, 59, 60]

Material	So (MPa)	S^ (MPa)	A (/sec)	ξ	m	n	Q/R (K)	a	ho (MPa)
SAC 305	1.3	39.4	500	7.1	0.3	.03	9000	1.5	5900
Sn63Pb37	12.4	14	4e6	1.5	0.3	.07	9400	1.3	1.38e3
	E (GPa)			ν			CTE (ppm/C)		
SAC 305	45 (25°C), 31.9 (100°C), 21.3 (200°C)			0.4			21.5		
Sn63Pb37	12.5(50°C), 6.9 (100°C), 5(200°C)			0.37			26 (50°C), 27.3 (100°C), 27.3 (200°C)		

SAC305 (96.5% tin, 3% silver, 0.5% copper) is used as the material for solder. Anand's viscoplastic constitutive model is used to describe the inelastic behavior of the Pb-free solder. This constitutive law follows the materials perspective that dislocation motion is the cause of both creep and plastic deformation, and combined them into inelastic strain. It unifies the creep and rate-independent plastic behavior of the solder by making use of a stress equation, a flow equation, and an evolution equation. The model needs no explicit yield condition and no loading/unloading criterion.

For the one-dimensional case (uniaxial loading), the stress equation is given by $\sigma = cs$; $c<1$, where s is the internal valuable and c is a function of strain rate and temperature expressed as

The total strain is expressed as,

$$\varepsilon_{ij} = \varepsilon_{ij}^{e} + \varepsilon_{ij}^{in} \qquad \text{Equation 3-1}$$

where ε_{ij}^{in} is the inelastic strain tensor.

The strain rate equation is represented by,

$$\frac{d\varepsilon_{in}}{dt} = A\left[\sin h\left(\xi\,\frac{\sigma}{s}\right)\right]^{\frac{1}{m}} \exp\left(-\frac{Q}{RT}\right) \qquad \text{Equation 3-2}$$

The rate of deformation resistance is given by,

$$\dot{s} = \left\{h_0(|B|)^{\alpha}\,\frac{B}{|B|}\right\}\frac{d\varepsilon_p}{dt} \qquad \text{Equation 3-3}$$

$$B = 1 - \frac{s}{s^*} \qquad \text{Equation 3-4}$$

$$s^* = \hat{s}\left[\frac{1}{A}\frac{d\varepsilon_p}{dt}\exp\left(-\frac{Q}{RT}\right)\right] \quad \text{Equation 3-5}$$

where $\frac{d\varepsilon_{in}}{dt}$ is the effective inelastic strain rate, σ is the effective true stress, s is the deformation resistance, T is the absolute temperature, A is pre-exponential factor, ξ is stress multiplier, m is strain rate sensitivity of stress, Q is activation energy, R is universal gas constant, h_0 is hardening/ softening constant, $\hat{s}$ is coefficient for deformation resistance saturation value, n is strain-rate sensitivity of saturation value, and α is strain-rate sensitivity of hardening or softening.

Table 3-3 Package Dimensions (full model)

Component	Thickness (µm)	Length (mm)	Width (mm)	Diameter (µm)	Height (µm)
Silicon Die	100	16	16		
Die Attach	10	-	-	10	100
PCB	1000	85	85		
Substrate	500	34	34	-	-
Solder Balls	-	-	-	500, 400, 350	275
Overmold	400	34	34	-	-

3.5 Design of Experiments

Feasibility of the proposed effective block technique has been demonstrated by analyzing 2 different solder alloys (SnPb and SAC 305) under reflow loading conditions for attachment to the PCB. The idea was to perform the following. 1) Demonstrate/Verify the scalability of the proposed compact modeling technique across different solder alloys

(SAC and SnPb), 2) to determine correlation between the ANAND's constants for the "effective block" and the solder volume fraction for SAC 305, 3) Qualitative validation of the simulation by comparing the trend of the mechanical response with existing literature and 4) scalability of the proposed technique across different loadings i.e., Reflow loading (monotonic) and Accelerated thermal cycling (cyclic).

Table 3-4 through Table 3-6 shows the different cases that have been studied to address the proposed work. Both the 'full array" and the "compact model" was simulated at each design point (case) and the response between the 2 models has been demonstrated. Global maximum warpage, solder joint von mises stress (sub-model), and the averaged plastic work in the corner solder joint was used as the validation parameters for the compact model. Different cases with varying solder diameters (different solder volumes) for each solder type – SAC 305 and Sn63Pb37 were analyzed to determine the correlation, if any between the respective "effective block" material constants and the solder volume fraction.

Table 3-4 Design of Experiments – Different solder footprints

Case	Solder mask opening (μm)	Solder volume (each ball) – mm3	Total solder volume 31x31 array (mm3)	Block Volume (mm3)
1	400	0.0538	51.7	318
2	300	0.0335	32.2	318
3	250	0.0250	24.2	318

Table 3-5 Mold Compounds with different Elastic Modulus (case 1 from Table 3-4)

Case	Young's Modulus (GPa)	Thermal Expansion Coefficient – ppm/C	Poisson's ratio ν	Total Solder Volume (31x31 Array) mm3
1	5	8, 32.5	0.3	51.7
2	25	8, 32.5	0.3	51.7
3	40	8, 32.5	0.3	51.7

Table 3-6 Different Solder Alloys – SAC 305 and SnPb

Case	Solder mask opening (μm)	Total solder volume (mm3)	Block Volume (mm3)	Solder Alloy
1	300	32.2	318	Sn63Pb37
2	300	32.2	318	SAC 305

3.6 Damage Prediction in BGAs (concept of the plastic work)

The solder joints subjected to monotonic or thermal cycling load tend to fail in the Low Cycle Fatigue (LCF) range due to thermo-mechanical failure. The fatigue life to failure generally falls between 100 to 10,000 thermal cycles. Depending on the test conditions and the fatigue damage parameters used, the selection of fatigue life models are made. The fatigue damage driving force parameters such as the plastic strain-range, creep strain range and inelastic strain energy density per cycle are a good correlation index to the BLR lifetime. [62, 63, 64].

The strain range-based fatigue approach deploys low cycle strain-controlled fatigue test method. The inelastic strain comprises of the plastic strain range and creep strain range. The plastic shear strain deformation is represented by the time-independent

plastic strain component, while the creep strain component contributes to the time-dependent inelastic strain included in the plastic shear strain ($\Delta\gamma_p$) component as shown in Figure 3-7. The energy-based fatigue model employs the cyclic stress–strain hysteresis loop to compute the elastic strain energy density (ΔW_e) and inelastic dissipated energy or plastic work per cycle (ΔW_p).

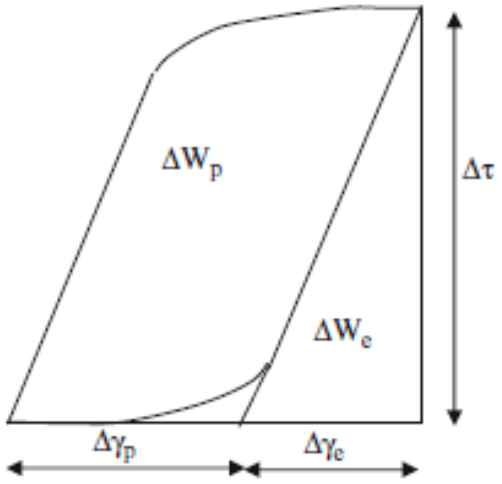


Figure 3-7 Cyclic stress-strain hysteresis loop

3.6.1 Energy Based Model

Darveaux [62] and a lot other groups have shown that the increment of inelastic strain energy density per thermal cycle can be used as a fatigue indicator. The inelastic strain energy density (inelastic strain energy per unit volume) is defined by

$$W^{in} = \int \sigma_{ij}\, d\varepsilon_{ij}^{in}$$

where σ_{ij} is the stress tensor and ε_{ij}^{in} in is the inelastic strain tensor. Since we are using Anand's constitutive law for the solder, the inelastic strain in this case is the viscoplastic strain.

Schubert proposed a fatigue model based on dissipated energy density during one thermal cycle and characteristic life. The model was proposed for PBGA's, CSP's and flip-chip packages for both SnPb and SnAgCu solder alloys (see Figure 3-8). In this work we will examine the model for SAC solder. The equation for the model is given by

$$N_f = 345\, W_{cr}^{(-1.02)}$$

where N_f is the characteristic life (cycles to 63.2% failure) and W_{cr} (unit in megapascal) is the strain energy density per cycle.

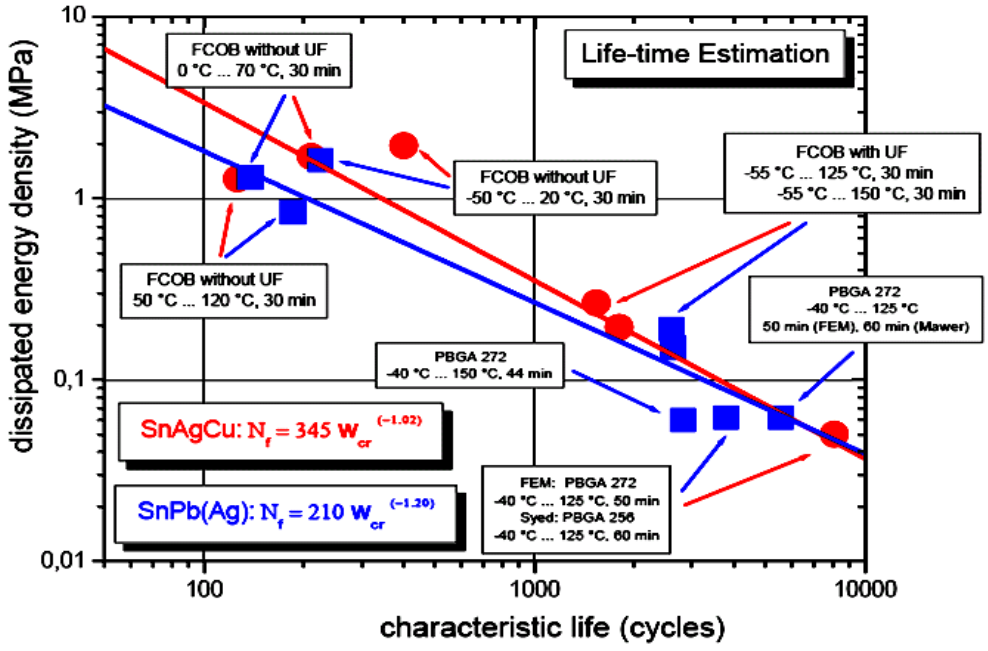


Figure 3-8 Schubert's energy based model for SAC and SnPb solder

Morrow's energy based model is used predict the low cycle fatigue life N_f in terms of inelastic strain energy density W_p (MPa) as shown below

$$N_f^n W_p = A$$

where n is the fatigue exponent and A is material ductility coefficient.

Syed [65] determined a life prediction model for CSP's and BGA's with SnAgCu solder material using strain energy density (or plastic work). The equation for the model is written below (see Figure 3-9)

$$N_f = 674.08\Delta W^{(-0.9229)}$$

The unit for plastic work (ΔW) is MPa or its equivalent MJ/m^3.

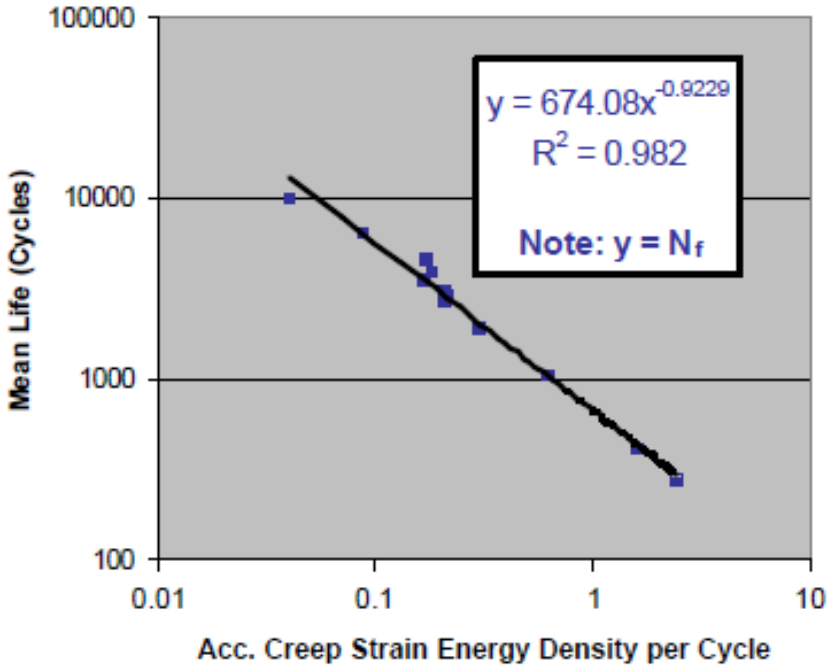


Figure 3-9 Syed's energy based model for CSPs and BGA

3.7 Results/Discussion

Entire modeling and analysis has been conducted on commercially available code ANSYS Workbench v14.5. Global maximum warpage, solder joint von mises stress, and solder joint equivalent strain in the corner solder joint were used as the parameters for validating the compact modeling technique. As expected, maximum damage was observed in the far corner joint in the full array simulation model. Mesh sensitivity analysis was performed on the global model to determine the critical mesh size beyond which there is negligible variation in the results. The global model at approximately 70000 elements was found to be the critical mesh-size of the model with mesh in-sensitive results. For all the design cases the global model was meshed at approximately 70000 elements to ensure mesh in-sensitivity in the results.

The loading ramp was 30°C/min from 200°C to RT for SAC and 170°C to RT for SnPb. Figure 3-10 shows the actual loading graph for the SAC305 alloy.

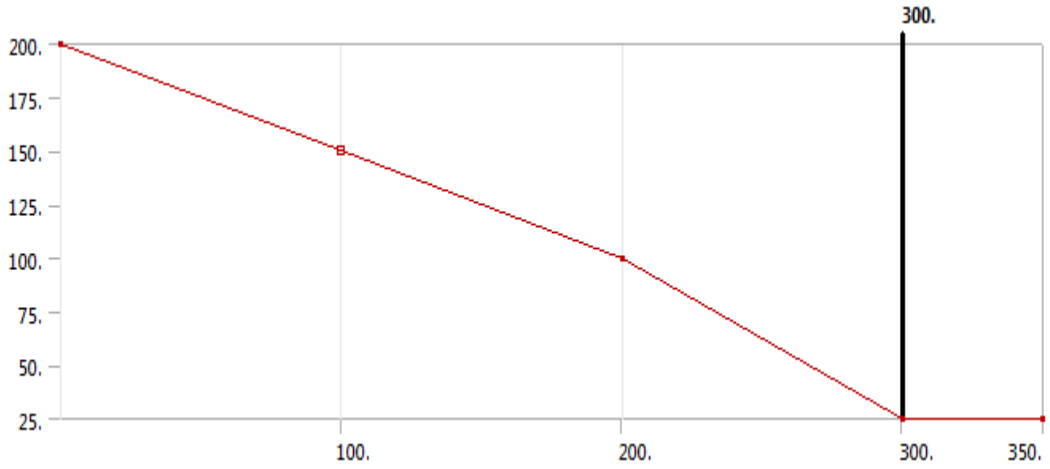


Figure 3-10 Reflow loading

3.7.1 Linear effective block

Initially, the effective block for the inner solder interconnect balls (visco-plastic material) was modeled as temperature dependent linear-elastic (for faster computational time). As expected, the results were significantly different thereby indicating that the assumption of the linear block was inaccurate. Solder is a viscoplastic material with time and temperature dependent inelastic behavior. The results below clearly demonstrate that the stress-strain magnitude (maximum von-mises stress and the total equivalent strain in the corner solder joint) between the array and the linear "effective block" compact model is quantitatively quite different from each other. The equivalent total strain in the corner solder joint observed with the "linear effective block" analysis is much smaller as compared to that of the array model. This is because when the effective block is considered linear, it does not undergo any inelastic (creep & plastic) deformation, thereby translating into reduced effect on the corner vulnerable region. Figure 3-11 through Figure 3-14 depict the inaccuracy of the linear "effective block" compact model when compared with the actual array model.

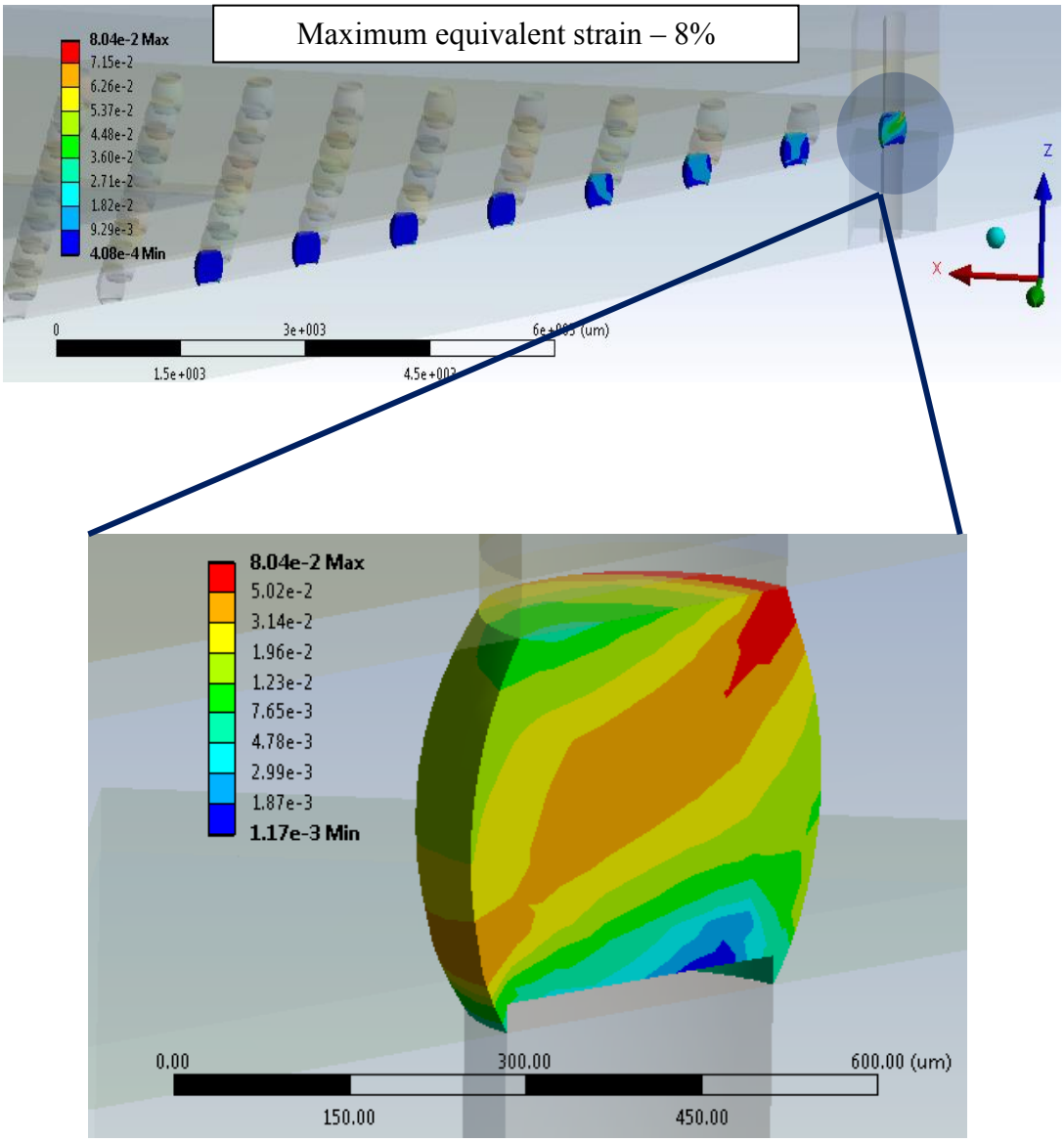


Figure 3-11 Corner joint strain – Array model

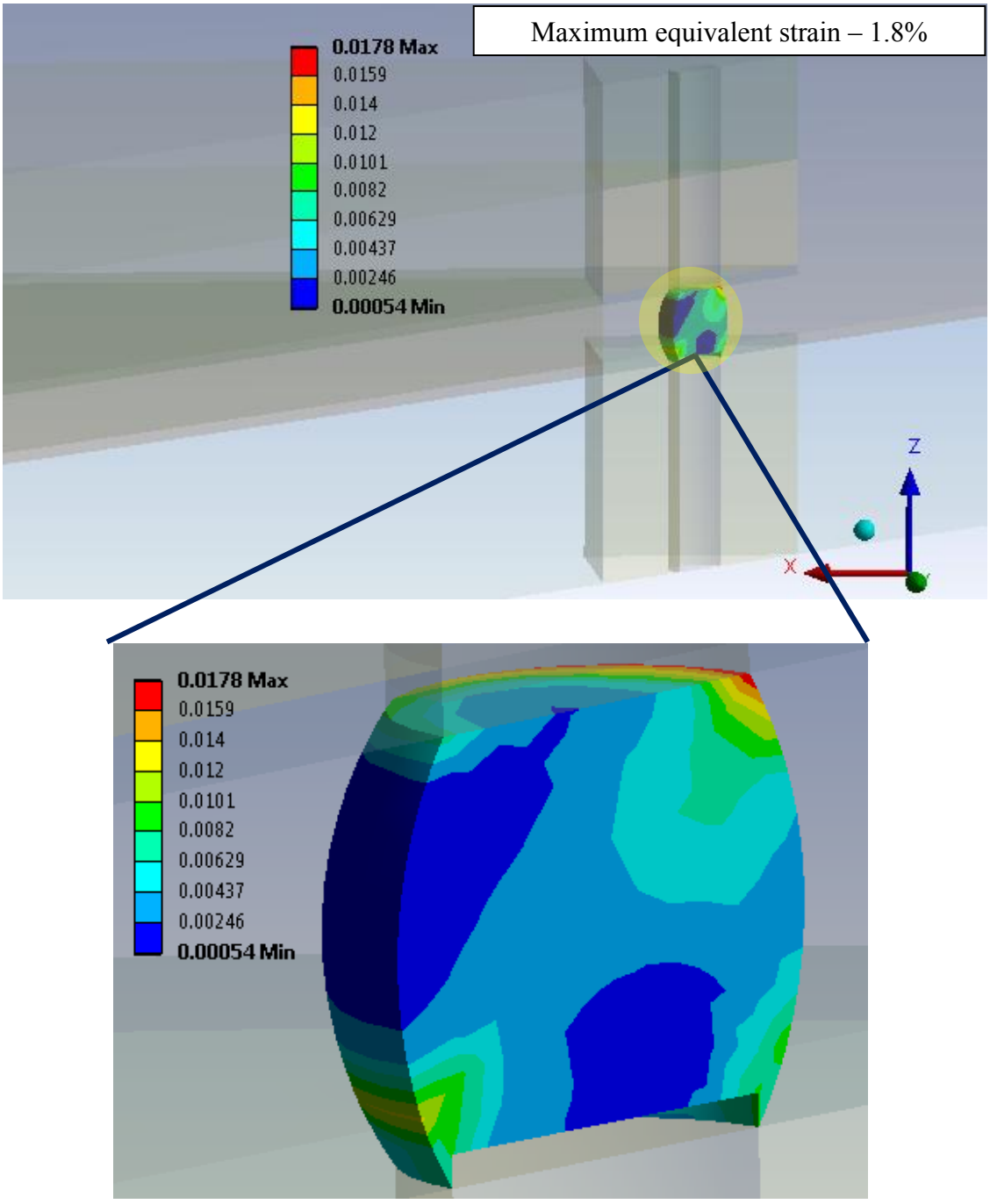


Figure 3-12 Corner joint strain – Linear "effective block" compact model

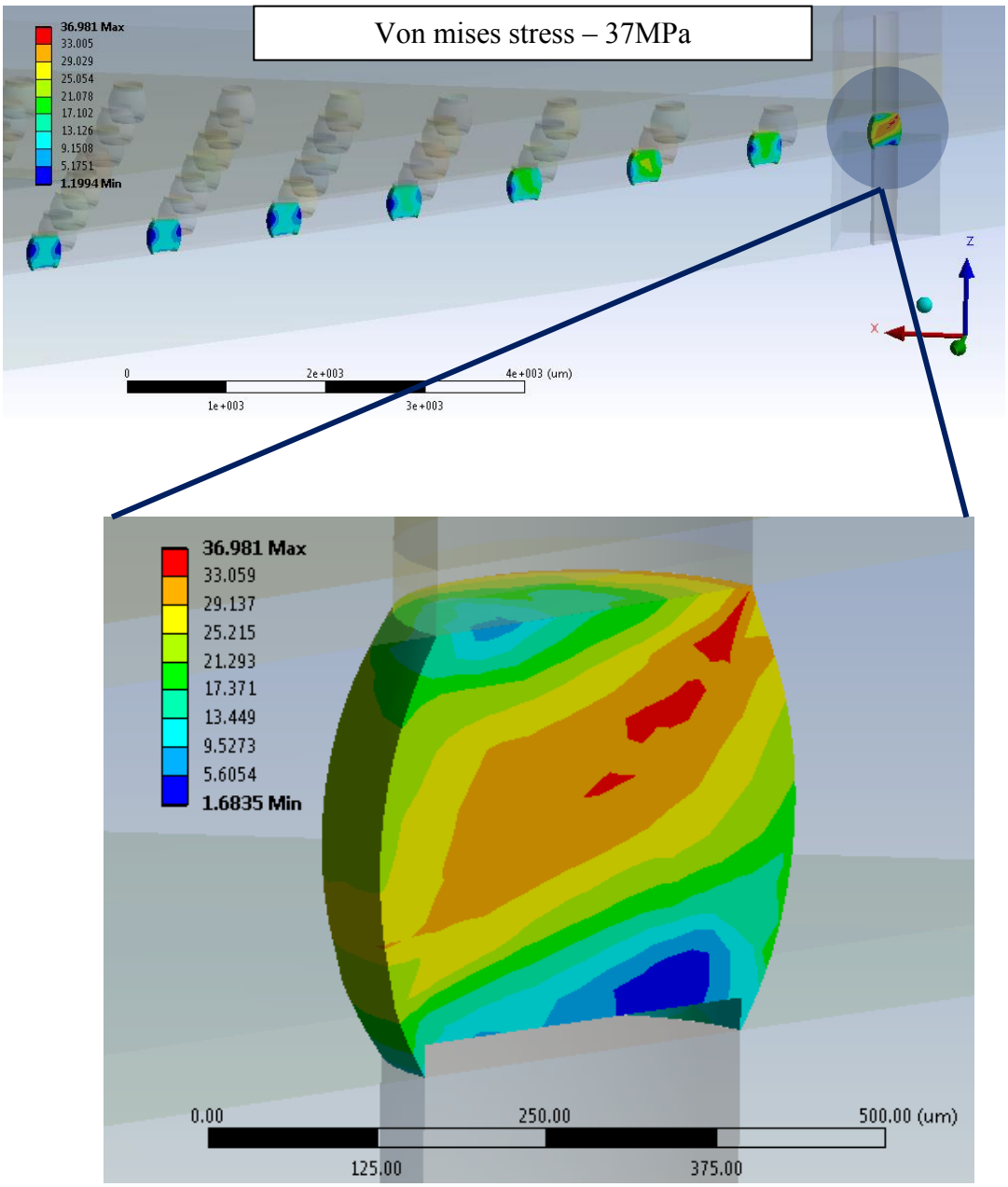


Figure 3-13 Corner joint von mises stress – array model

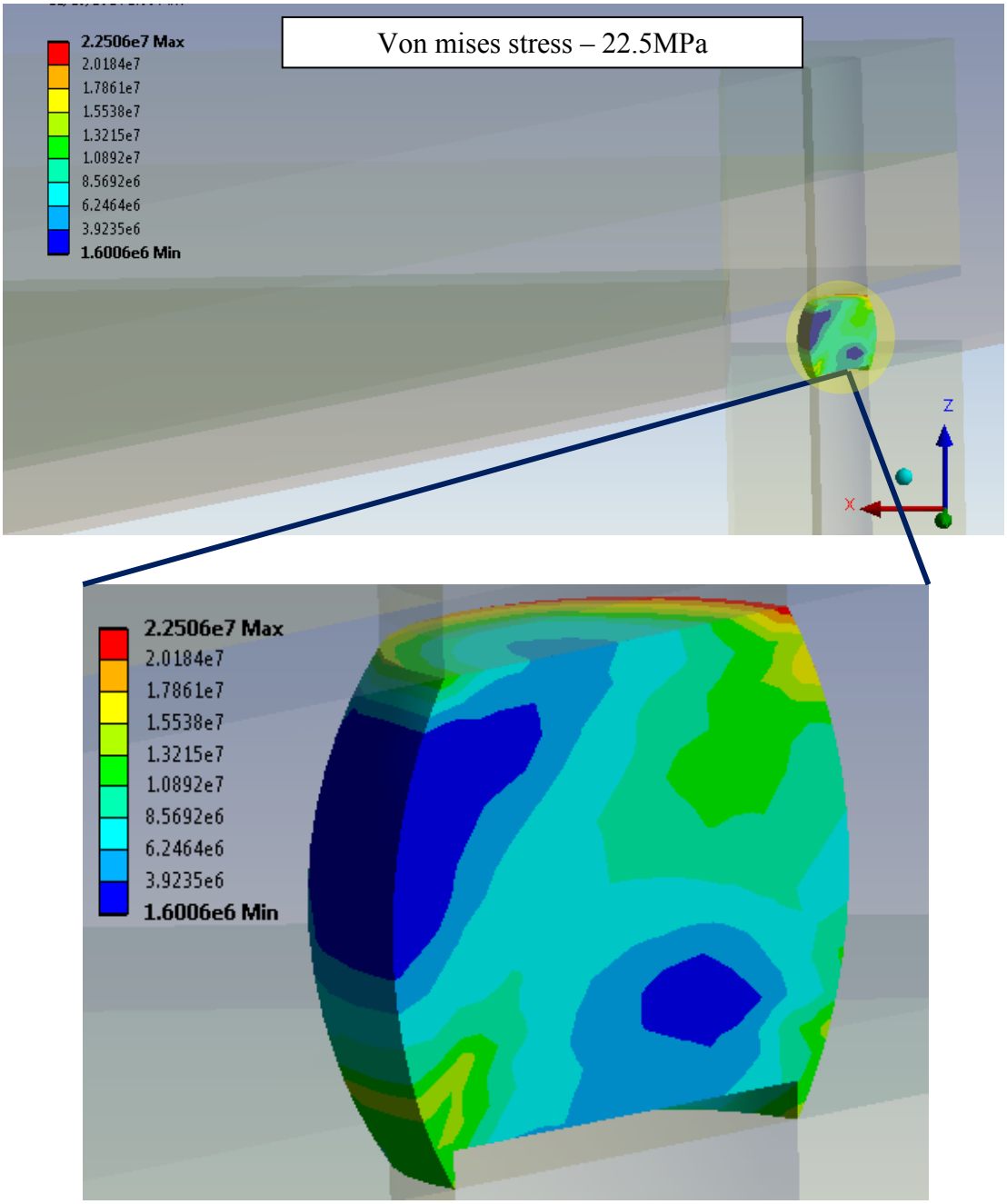


Figure 3-14 Corner joint von mises stress – linear “effective block” compact model

Table 3-7 “Linear effective block” vs “Array” configuration

Case	Corner joint stress - MPa			Corner joint strain - %		
	Array model	Compact model (linear block)	% difference	Array model	Compact model (linear block)	% difference
2	37	22.5	-39%	8	1.8	80

As shown in Table 3-7, there is a huge error in the thermo-mechanical predictions by using a linear “effective block” approach. Therefore, the effective block for the inner region of the solder balls is modeled as viscoplastic but the challenge was to determine the material model parameters for which the mechanical response between the array and the compact model was similar. 2 constants (based on their definition) from the actual viscoplastic ANAND’s model were tuned in such a way that the effect of the “effective block” on the corner region was same as that of the inner balls. s_o (coefficient of initial deformation resistance) and S^ (coefficient of deformation resistance) were tuned in a such a way that the effect of the proposed block on the thermo-mechanical response is similar to that of the array configuration. The next section demonstrates the validation of the compact model with the array model for different solder footprints.

3.7.2 Validation of SAC 305 Compact Model (non-linear block)

In all, 3 cases with varying solder footprints (refer to Table 3-4) have been studied to demonstrate the feasibility of the compact modeling technique and also develop a correlation of the modified ANAND's viscoplastic material constants for the effective block with the solder volume fraction. Figure 3-15 through Figure 3-24 show the comparison between the SAC 305 "full array" model and the "effective block" compact model for global warpage, solder joint von-Mises stress, and the solder joint total strain for case 1 ($51.7mm^3$) and case 3 ($24.2mm^3$). The results show that the non-linear material model for the effective block is the way to go as it poses a similar effect on the corner region as that done by the actual array configuration. In non-temperature loadings (pure mechanical type of loadings), it may be possible to get away with the linear block due to the fact that the solder would not creep as much as it does at higher temperatures. The inelastic strain (creep + plastic) in the solder at room temperature would be lesser in a room temperature analysis but it all depends on the extent of loading because the solder material has a very small linear zone and the majority of strain observed in such materials comes from the inelastic region.

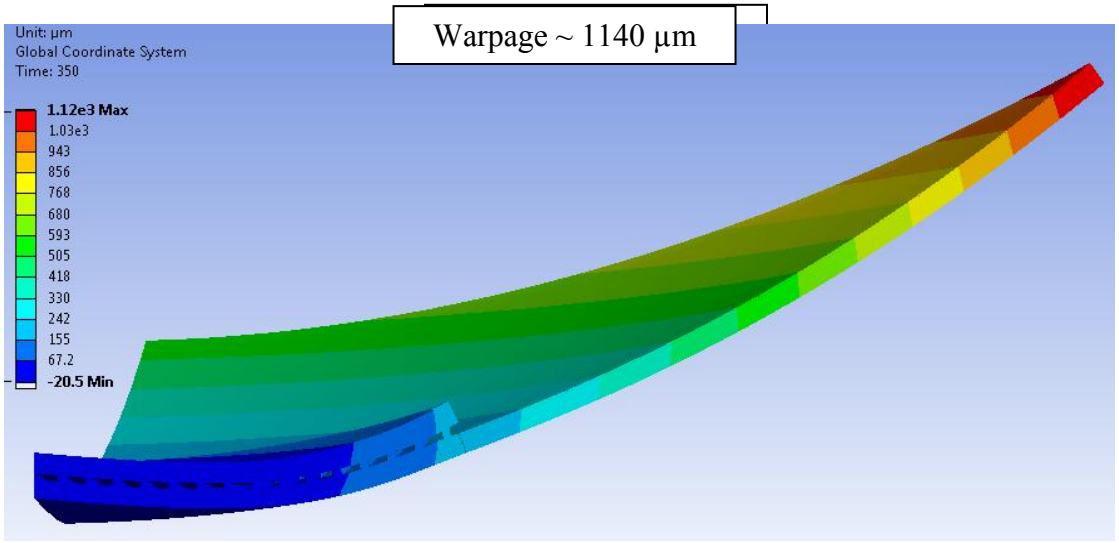


Figure 3-16 Case 1 – Array model warpage (SAC 305)

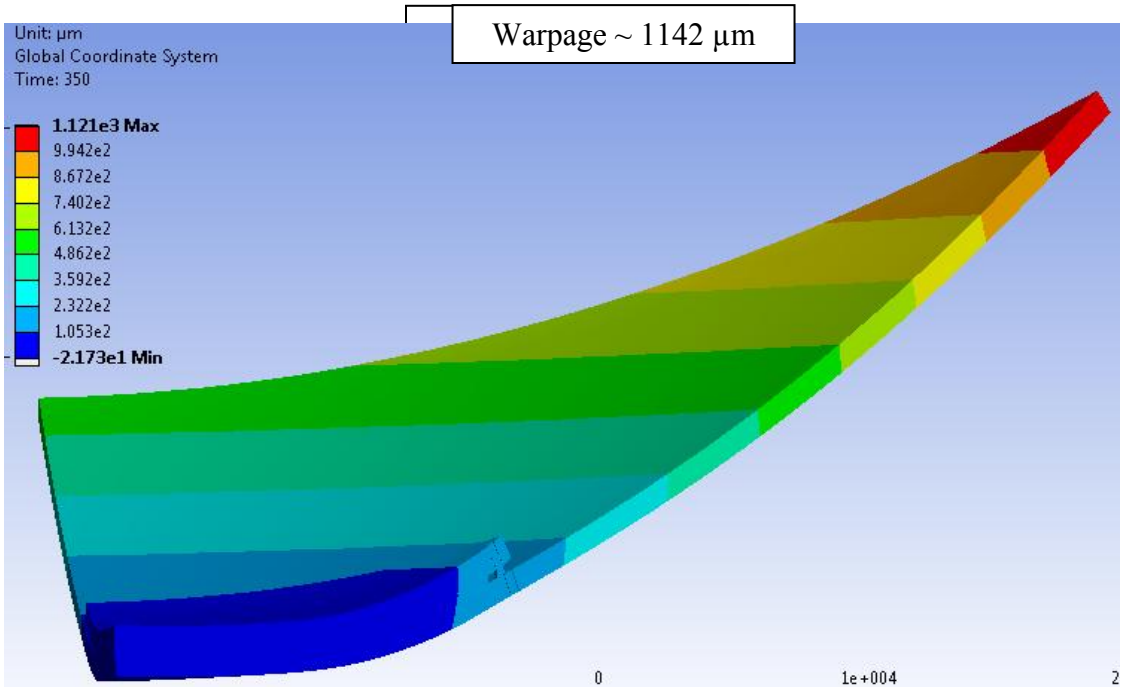


Figure 3-15 Case 1 – Compact model warpage (SAC 305)

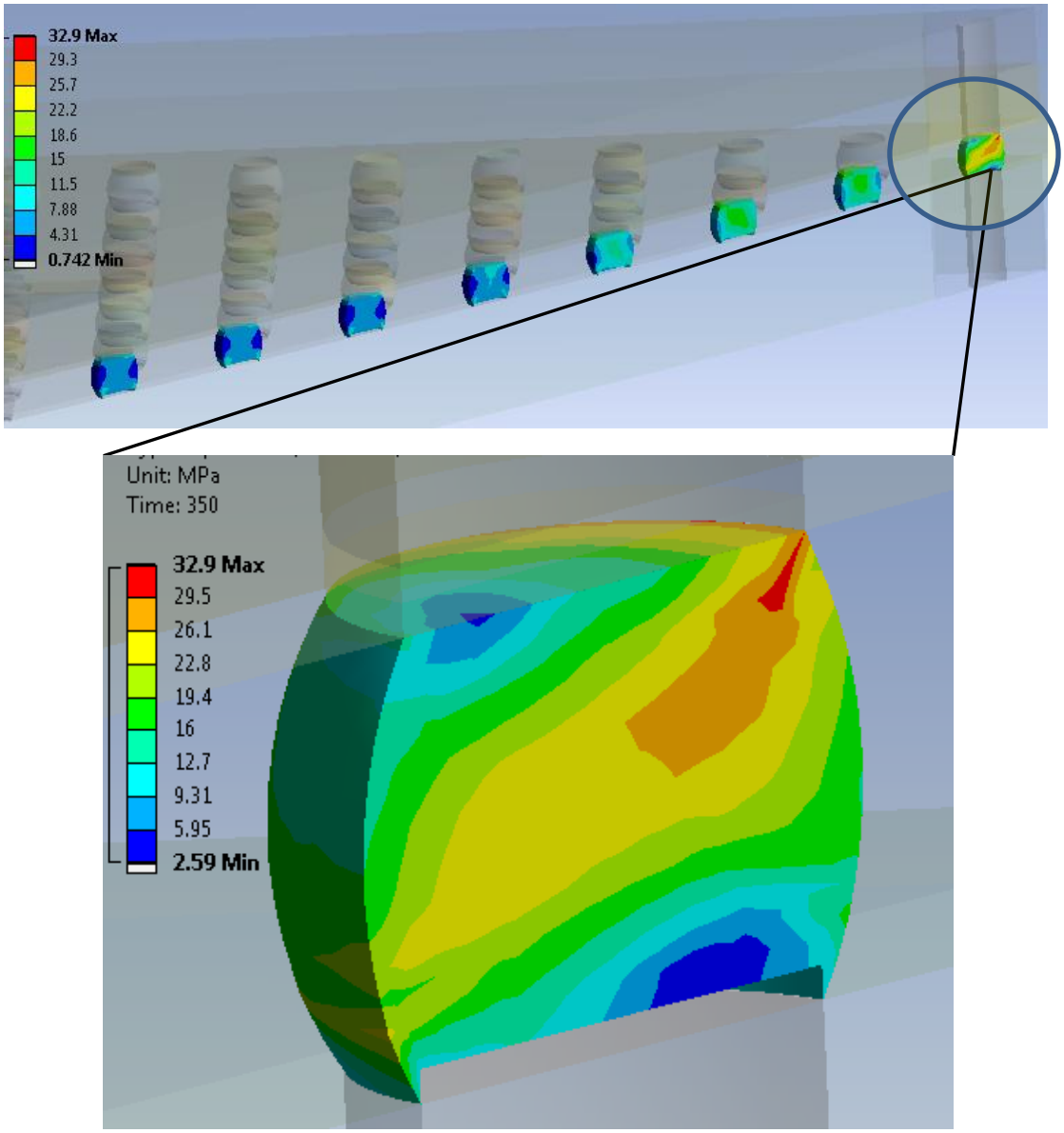


Figure 3-17 Case 1 – corner joint von-Mises stress (Array Model SAC 305)

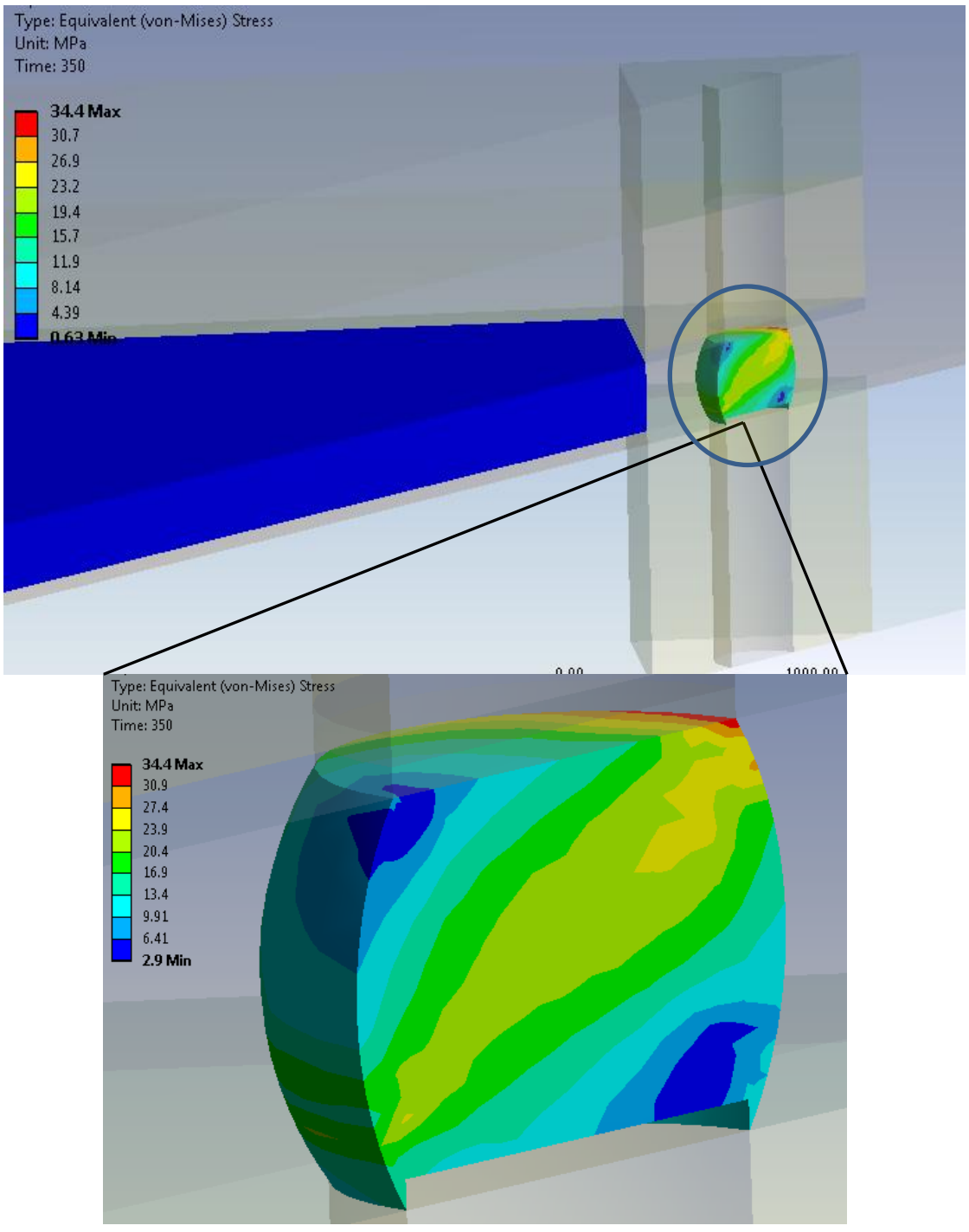


Figure 3-18 Case 1 - Corner joint von Mises stress (compact model - SAC 305)

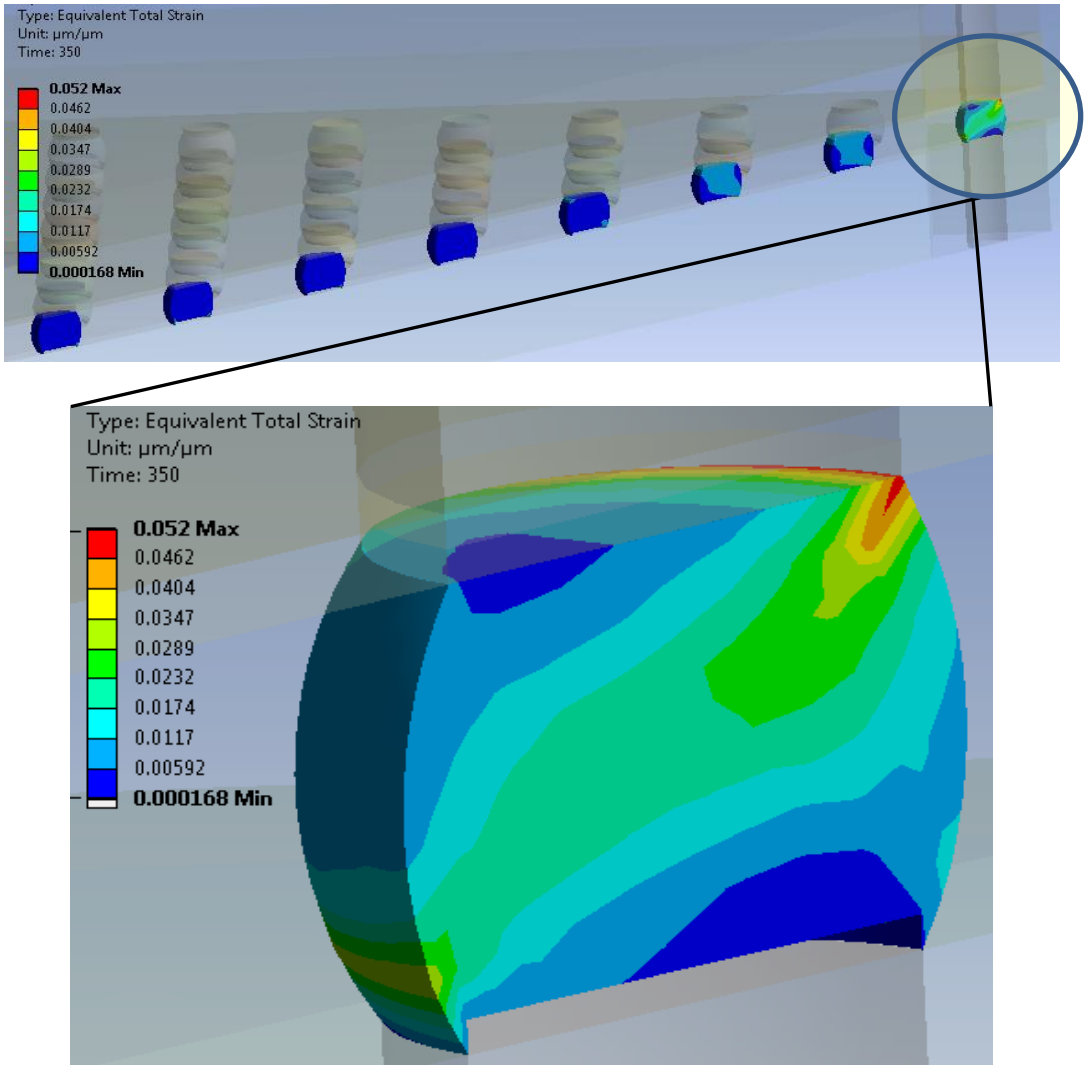


Figure 3-19 Case 1 – Equivalent total strain (Array model SAC 305)

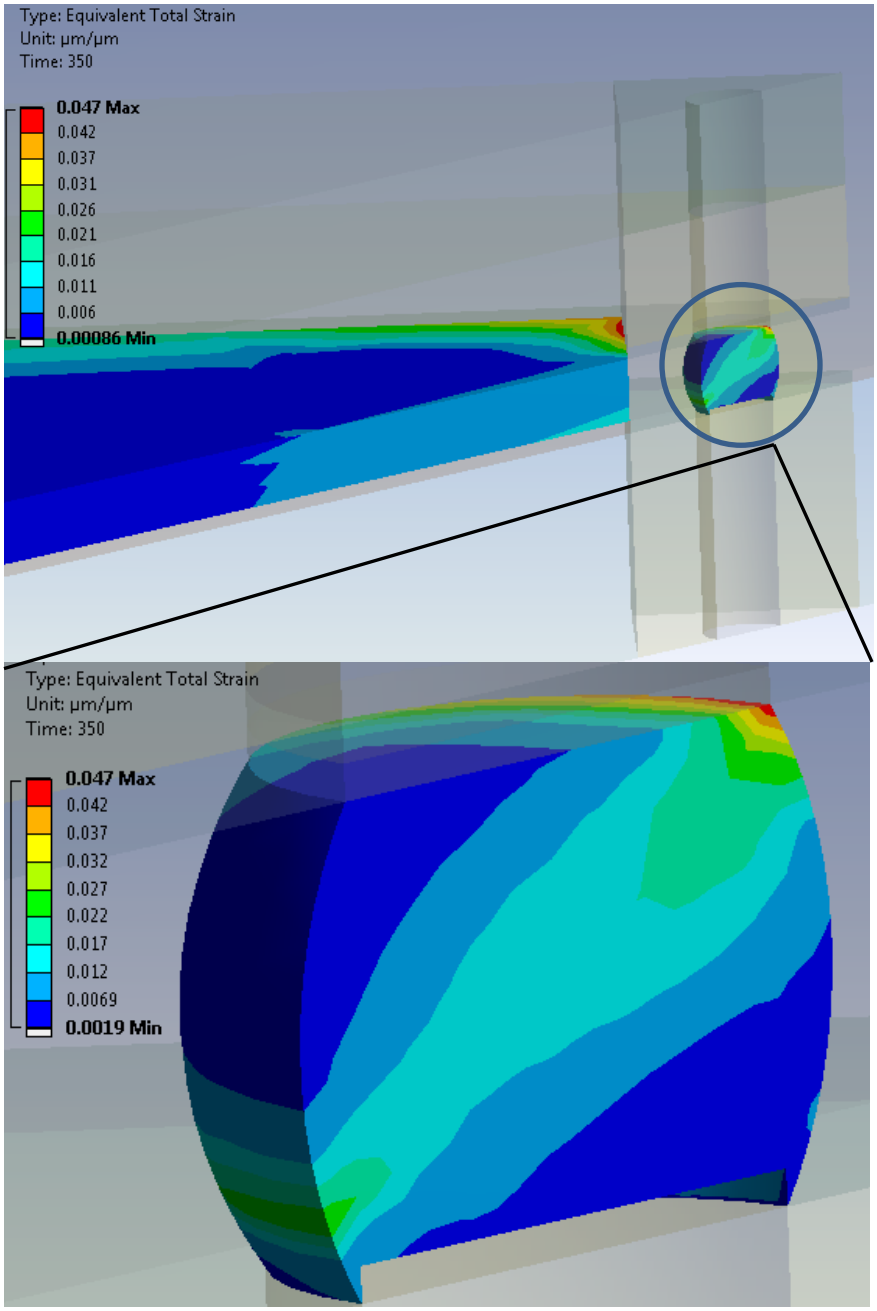


Figure 3-20 Case 1 – Equivalent total strain (compact model SAC 305)

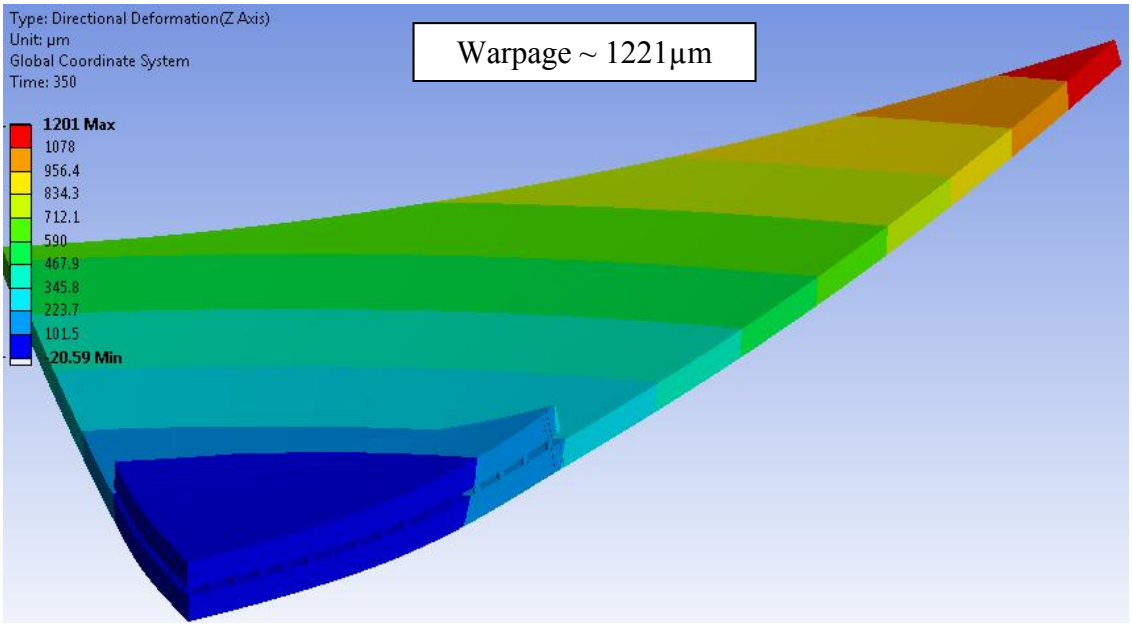


Figure 3-21 Case 3 – Array model warpage (SAC 305)

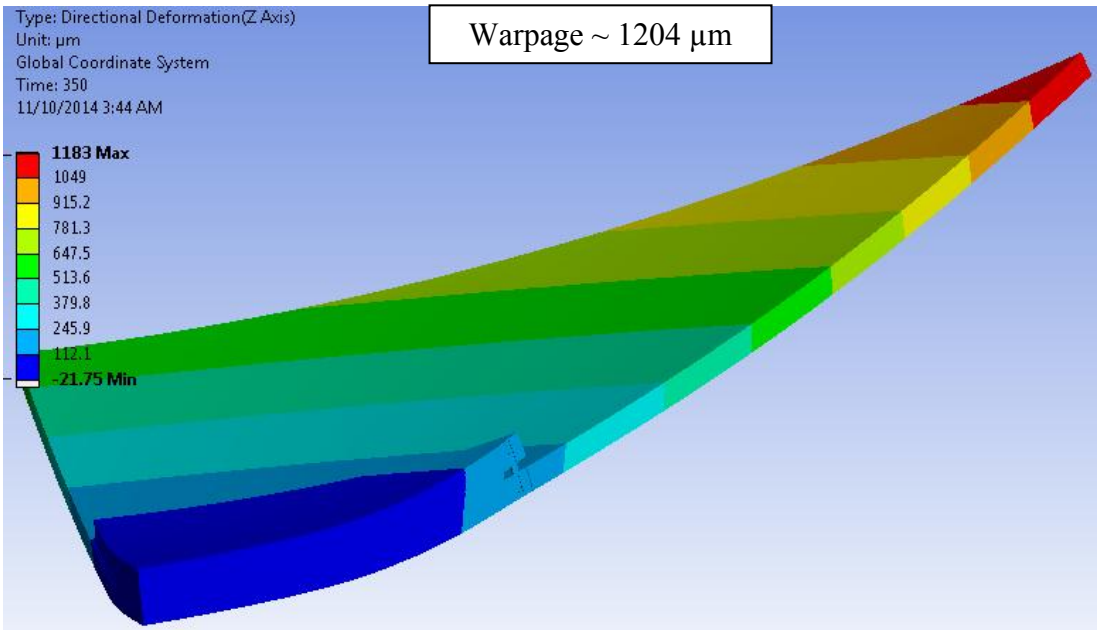


Figure 3-22 Case 3- Compact model warpage (SAC 305)

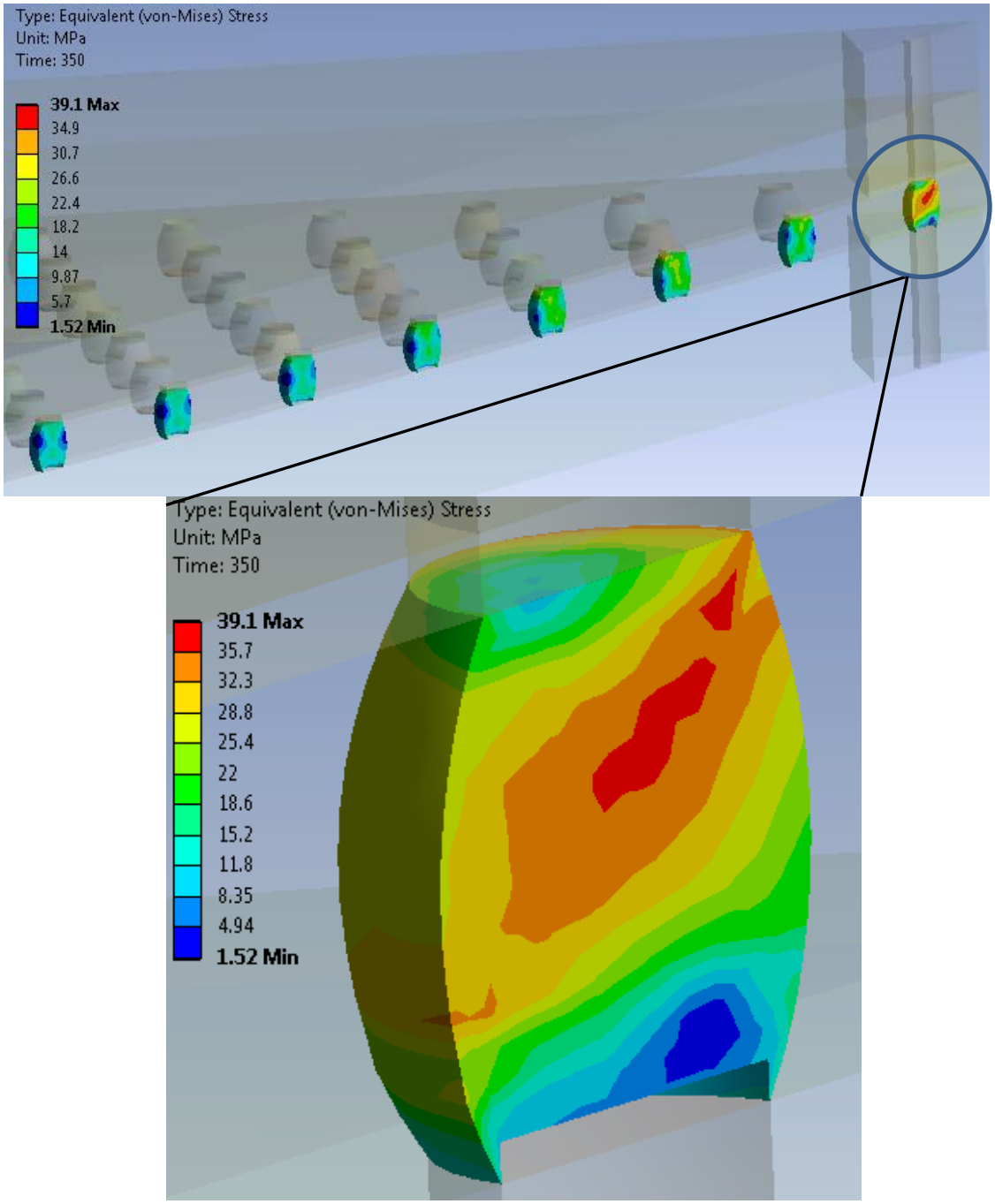


Figure 3-21 Case 3 - Corner joint von-Mises stress (Array model SAC 305)

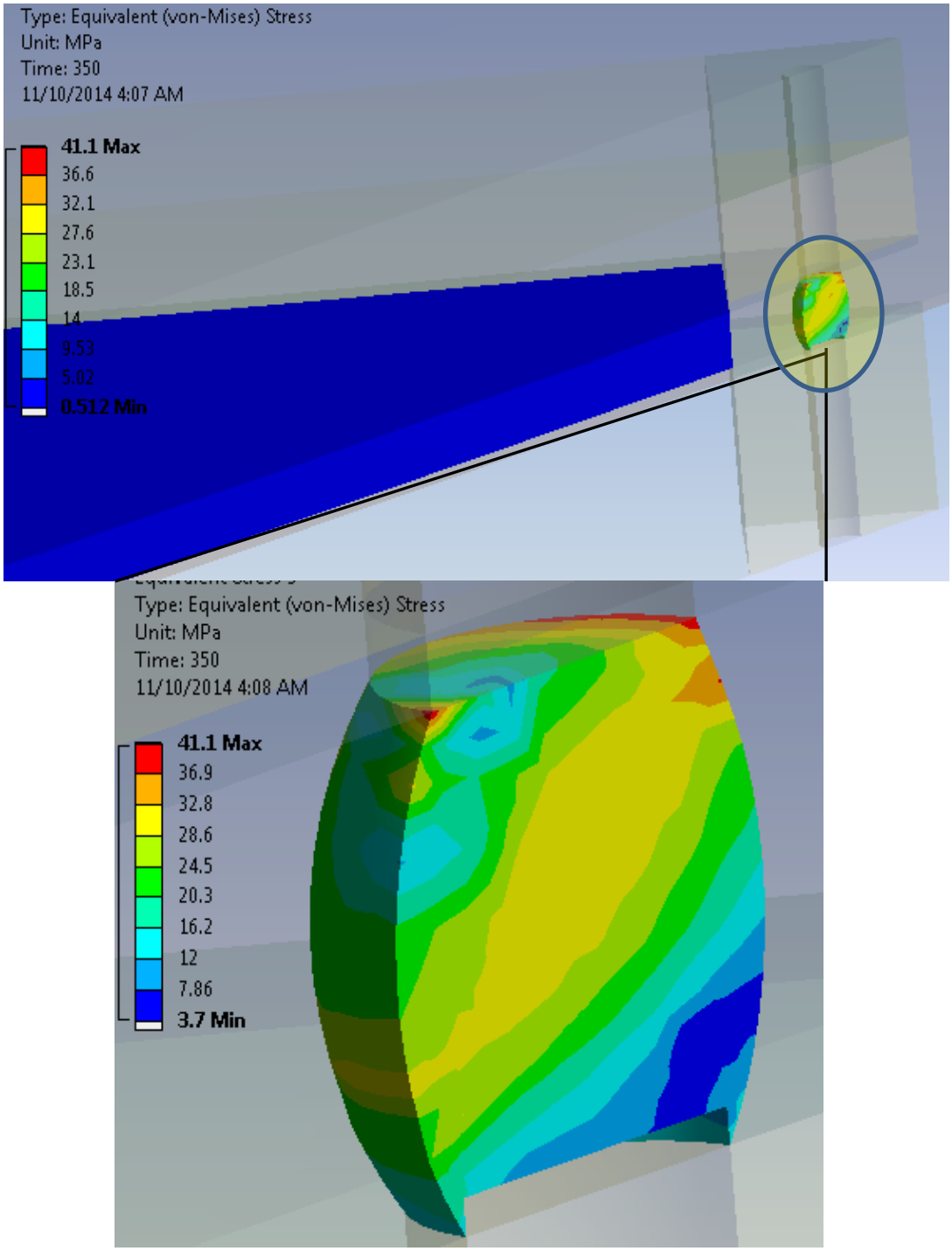


Figure 3-22 Case 3- corner joint von-Mises stress (Compact model SAC 305)

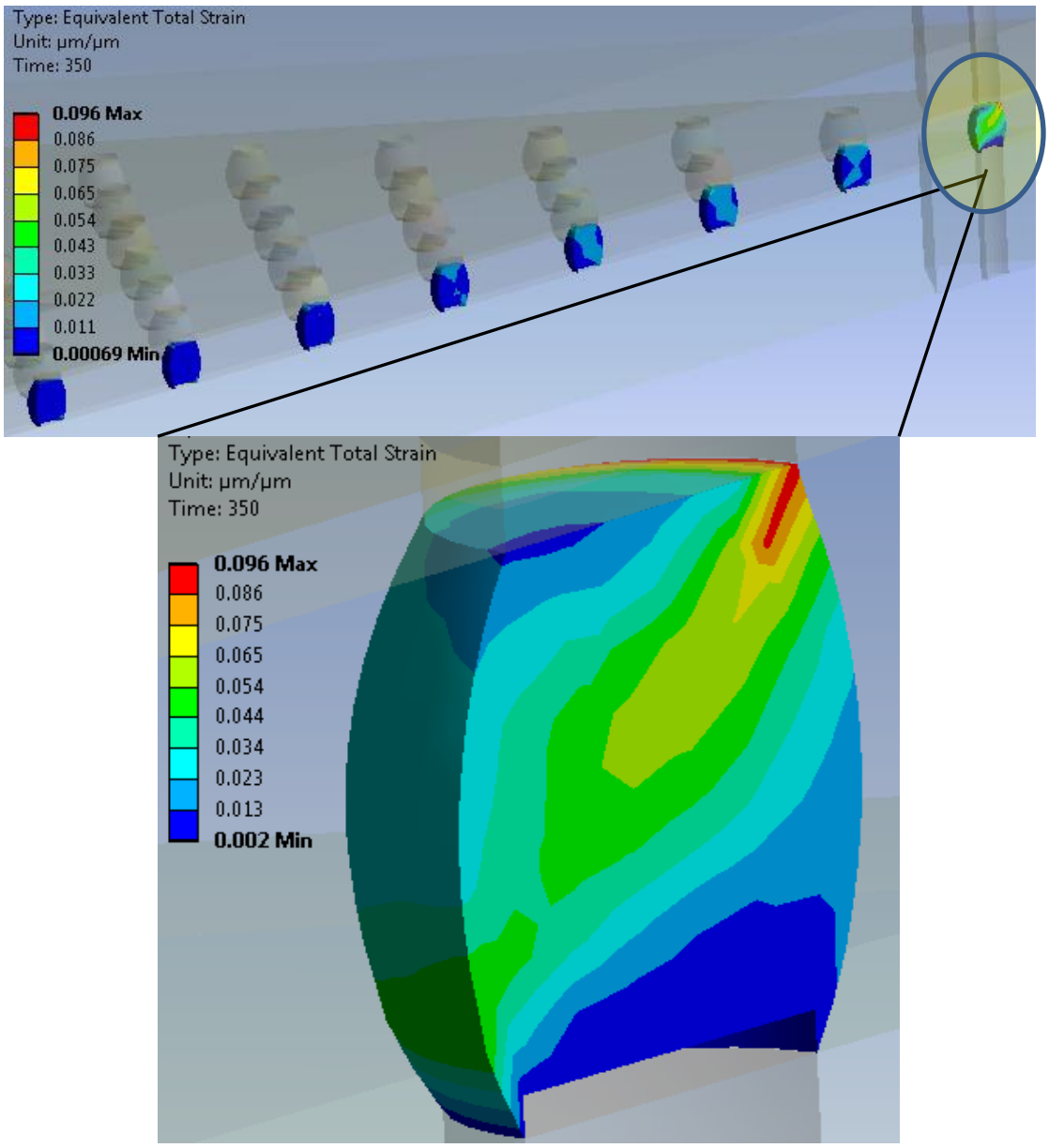


Figure 3-23 Case 3 – Corner joint equivalent total strain (Array model SAC 305)

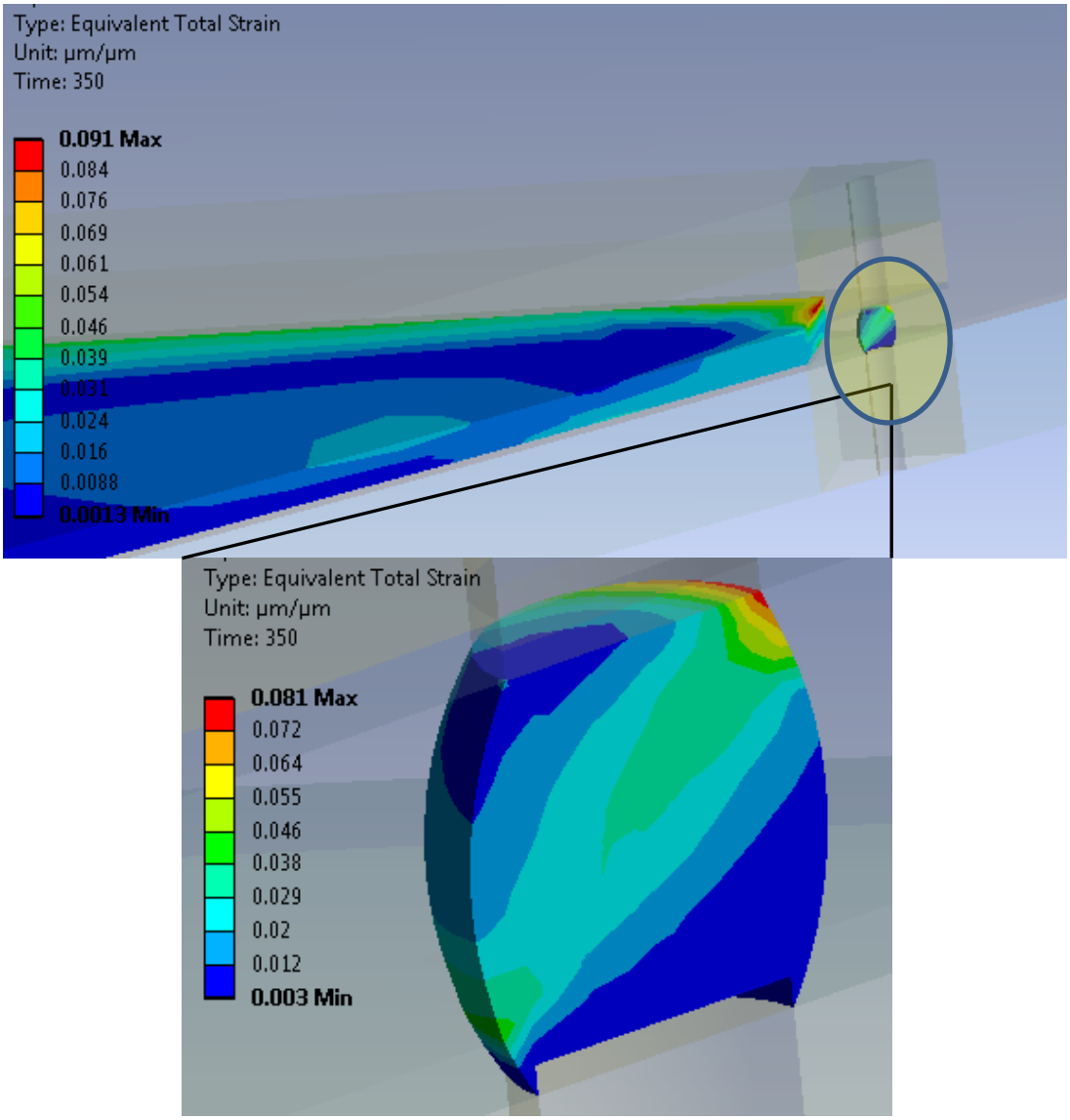


Figure 3-24 Case 3 – Corner joint strain (compact model SAC 305)

It is clearly evident from the above contour plots that there is a good agreement between the full array model and the compact model mechanical response. The maximum difference between the compact and the array model was seen for case 3 (equivalent strain ~ 16% error). Thermo-mechanical prediction by the compact model is consistent

with the array model thereby making this compact modeling methodology a viable technique to alleviate the torturous computational time and lead to faster development. It is to be noted that the compact model showed up to 15x computational time reduction as compared with the array counterpart. This huge computational time reduction is due to the element size difference between the array and the compact model (70000 elements in array model as compared to 3000 elements in the compact model). Since most of the element reduction came from the solder balls (solder balls changed to an effective block), which is visco-plastic, the effect on the computational time was huge.

The output response from both the array and the compact model is in line with literature (qualitatively), the damage in the solder joint increases as the mask opening size reduces (as shown below in Table 3-9 and Figure 3-25 to Figure 3-27) [61].

Table 3-8 Summary of results – Different solder footprint

Case	Solder Volume (mm3)	Warpage (um)			Maximum von-Mises stress Corner joint - MPa			Maximum Equivalent Strain - % Corner joint		
		Array model	Compact model	% Error	Array model	Compact model	% Error	Array model	Compact model	% Error
1	51.7	1138	1142	-	32.9	34.3	4.2	5.2	4.7	9
2	32.2	1200	1194	0.5	37	39.5	6.8	8.04	7.2	10
3	24.2	1221	1203	1.5	39.1	41.1	6	9.6	8.1	16

Table 3-9 Effect of solder mask opening on the mechanical response

Case	Solder mask opening (µm)	Warpage (um)			Maximum von-Mises stress Corner joint - MPa			Maximum Equivalent Strain - % Corner joint		
		Array model	Compact model	% Error	Array model	Compact model	% Error	Array model	Compact model	% Error
1	400	1138	1142	-	32.9	34.3	4.2	5.2	4.7	9
2	300	1200	1194	0.5	37	39.5	6.8	8.04	7.2	10
3	250	1221	1203	1.5	39.1	41.1	6	9.6	8.1	16

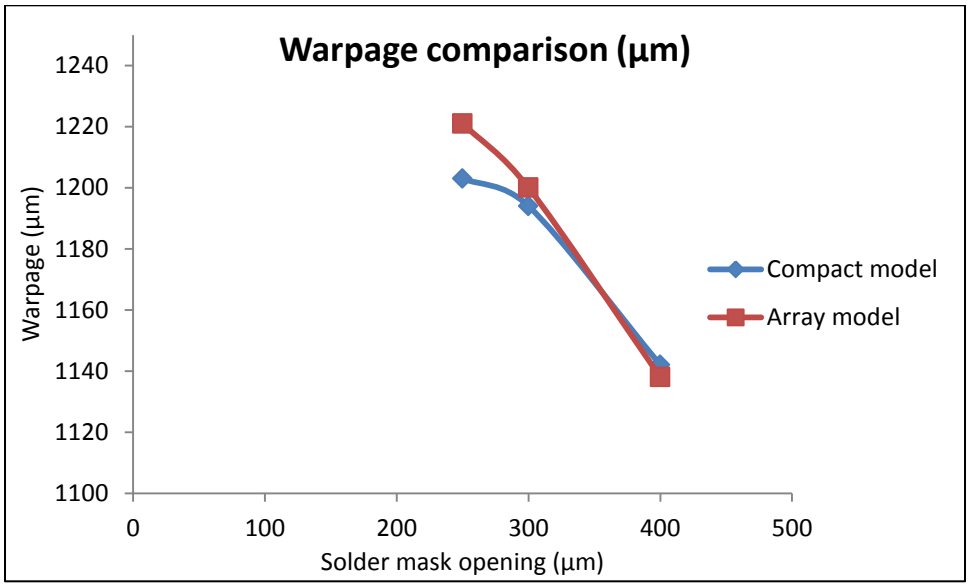


Figure 3-25 Warpage variation with solder mask opening

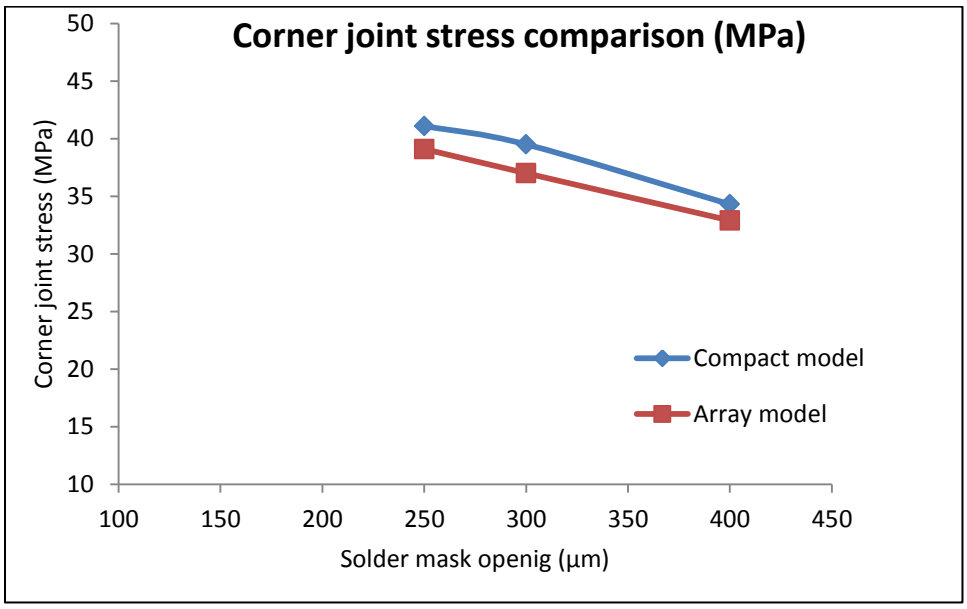


Figure 3-26 Corner joint stress variation with mask opening

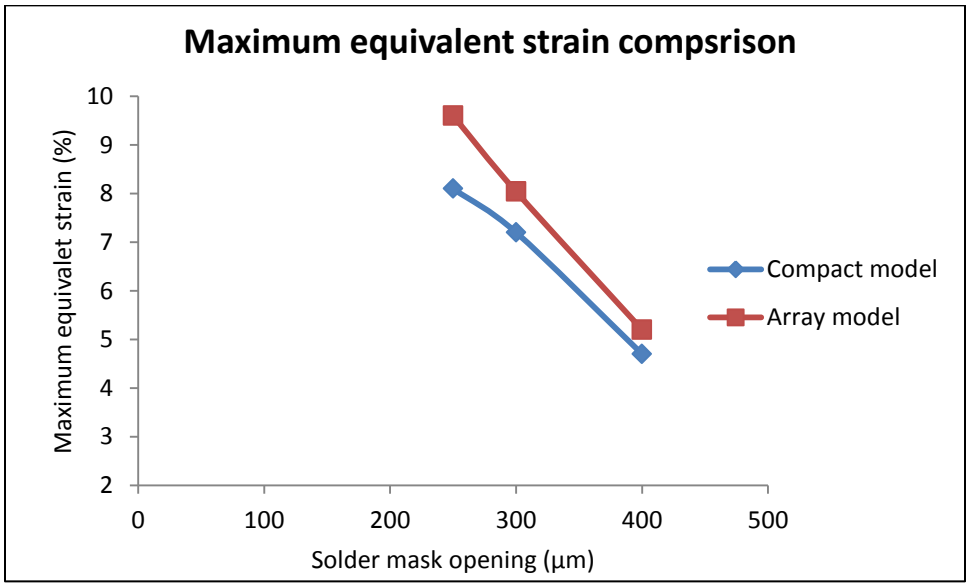


Figure 3-27 Corner joint strain variation with mask opening

Table 3-10 Modified viscoplastic model constants for the effective block

Case	Solder Volume – Vs (mm3)	Effective Block Volume – Vb (mm3)	Actual ANAND's Model Constant (so) MPa	Actual ANAND's Model Constant (S^) - MPa	Compact model Modified Constant (so) MPa	Compact model Modified Constant (S^) - MPa
1	51.7	318	1.3	39.4	0.25	1.7
2	32.2	318	1.3	39.4	0.2	1.07
3	24.2	318	1.3	39.4	0.175	1.0165

Based on the comparative results between the array and the compact model, a quadratic relation is developed that correlates the modified ANAND's constants with the solder volume fraction for SAC 305 alloy. This relation will lead to determining the new constants instantaneously without having the need to perform tuning across different solder footprints (solder diameter, mask opening, array size, and solder height), different material selection DOEs, and so on. However, there is a limitation with this technique - this quadratic relation is valid for a particular solder alloy and a new relation will have to developed whenever the solder alloy changes. Figure 3-28 through Figure 3-30 show the visual summary of the excellent correlation between the array and the compact model.

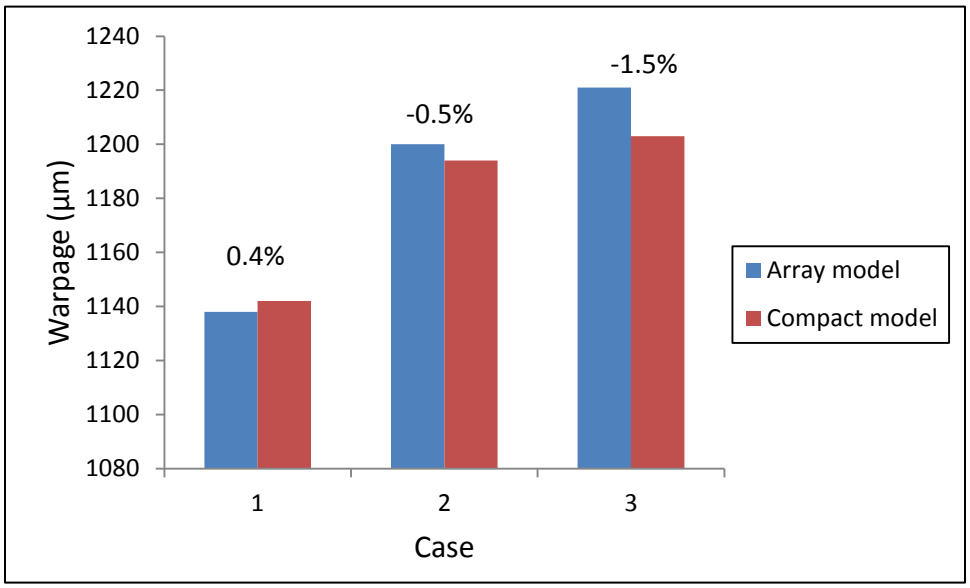


Figure 3-28 Warpage comparison between "array" and "compact" model

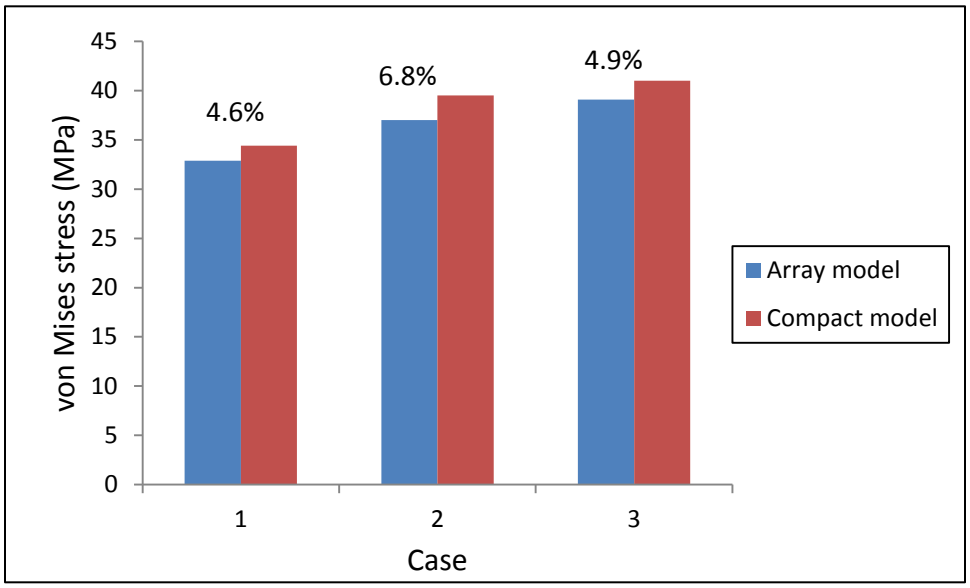


Figure 3-29 Stress comparison between "array" and "compact" model

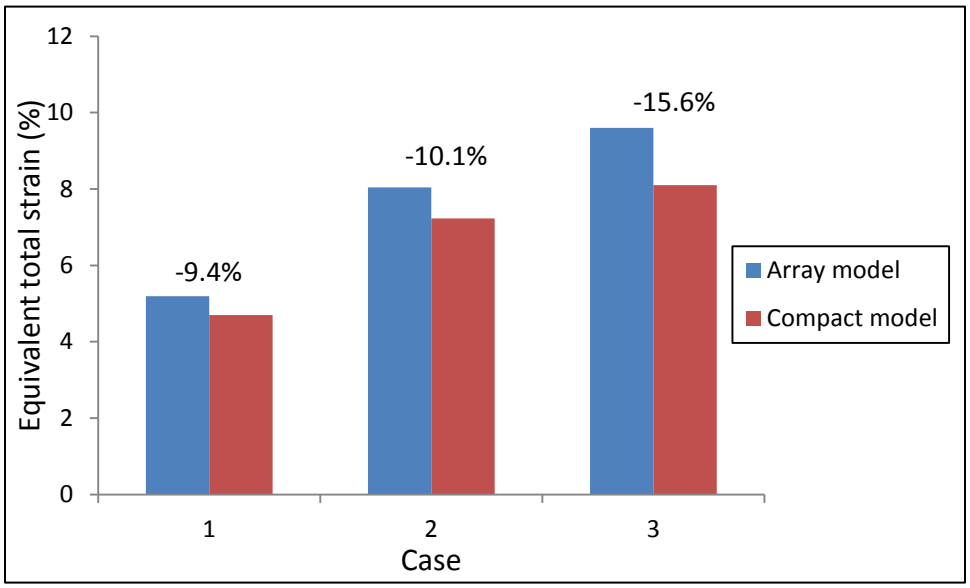


Figure 3-30 Strain comparison between "array and "compact" model

Interestingly, it is to be noted (from Table 3-10) that as the solder footprint is changed (such a way that the solder volume is decreased) there is a particular direction the new material constants follow. The magnitude of the constants (s_o and S^) is much less as compared with the actual ANAND's model for the solder ball (see Table 3-10) owing to that fact that the effective block's material model should compensate for the material between the balls (i.e. in this case its air). A non-dimensional relation (quadratic in nature) is formulated through this work that correlates the new constants with the ratio of solder volume to the effective block volume (see Figure 3-31and Figure 3-32).

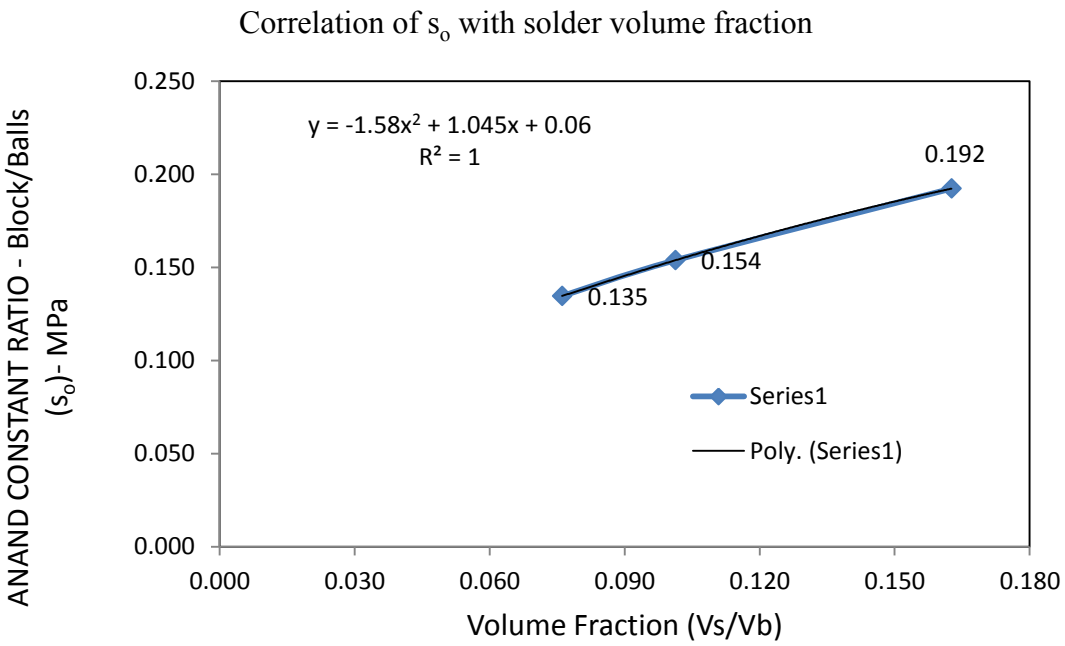


Figure 3-31 Correlation model for s_o

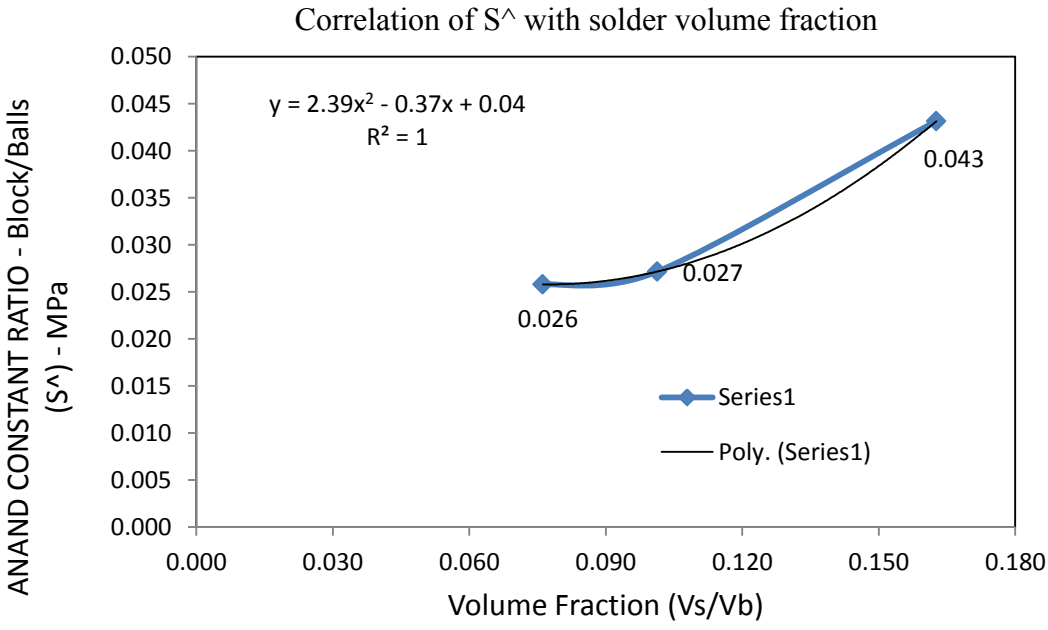


Figure 3-32 Correlation model for S^

3.7.3 Qualitative validation of the compact model with literature (effect of mold compound elastic modulus on the corner joint response)

In order to demonstrate the credibility of the numerical model (both array and compact model) and also show that the proposed constants accurately predict the mechanical response of the package with different packaging materials, a parametric study was performed with different overmold compounds (varying Young's modulus) and the output response was qualitatively compared with the existing literature . This study shows that the tuned material constants for the effective solder block are only dependent on the solder type and the volume fraction of the solder. In other words, if the modified constants for the effective solder block have been determined for a particular solder alloy and the model has been validated with the full array at any one design point, the subsequent design points in the DOE can be solved using the compact model without compromising the accuracy. As per the literature [61], stiffer mold compound increases the damage to the solder joint. 3 different mold compound materials were analyzed (E =5, 25, 40GPa) and it was seen that corner joint stress-strain magnitude in both the array and the compact model substantially increased with the stiffer mold, thereby further strengthening the credibility of the simulation work and more importantly showing that the compact model is a viable and efficient technique for such analysis.

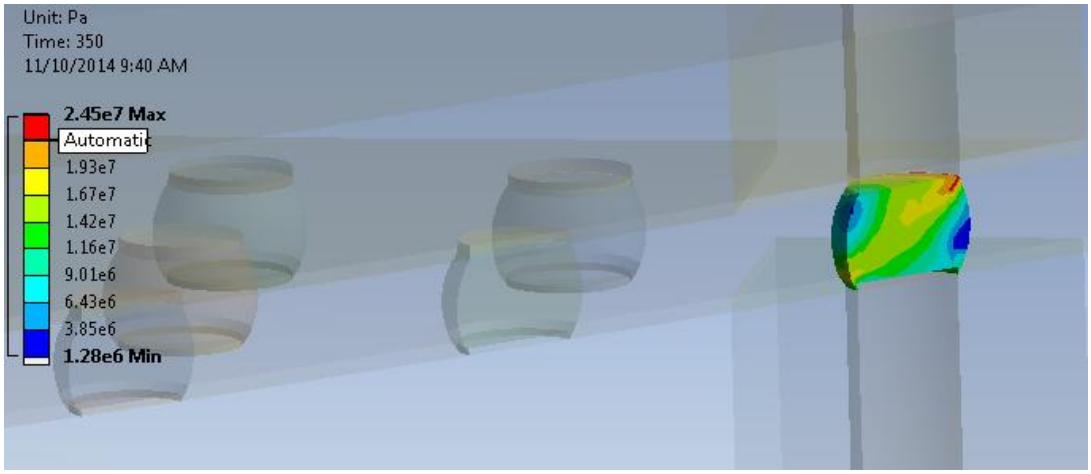


Figure 3-33 Array model - von-Mises stress (mold E = 5GPa)

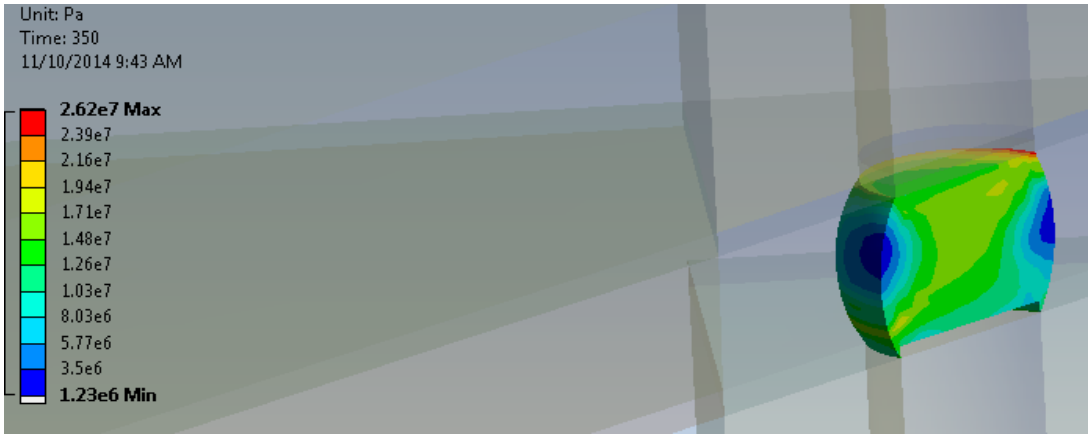


Figure 3-34 Compact model – von-Mises stress (mold E = 5GPa)

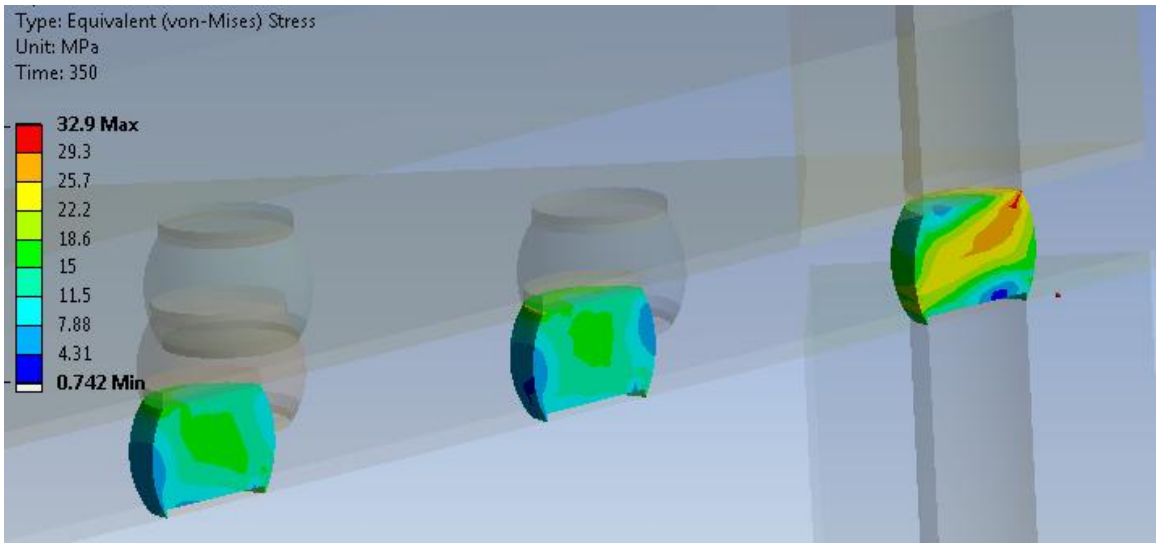


Figure 3-37 Array model – von mises stress (mold E = 25GPa)

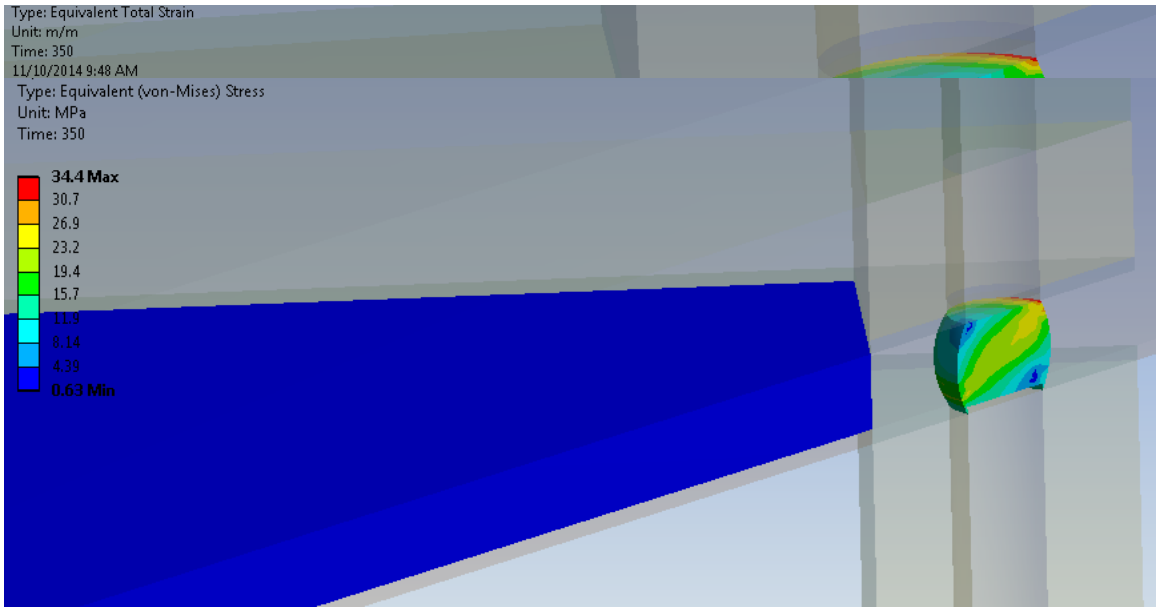


Figure 3-38 Compact model – von Mises stress (mold E = 25GPa)

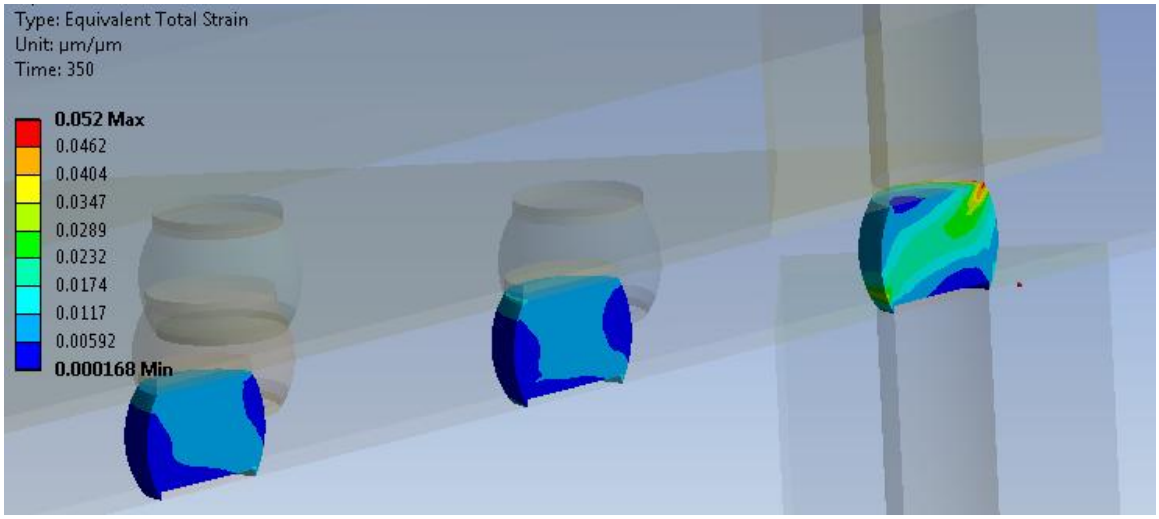


Figure 3-39 Array model – Equivalent total strain (mold E =25GPa)

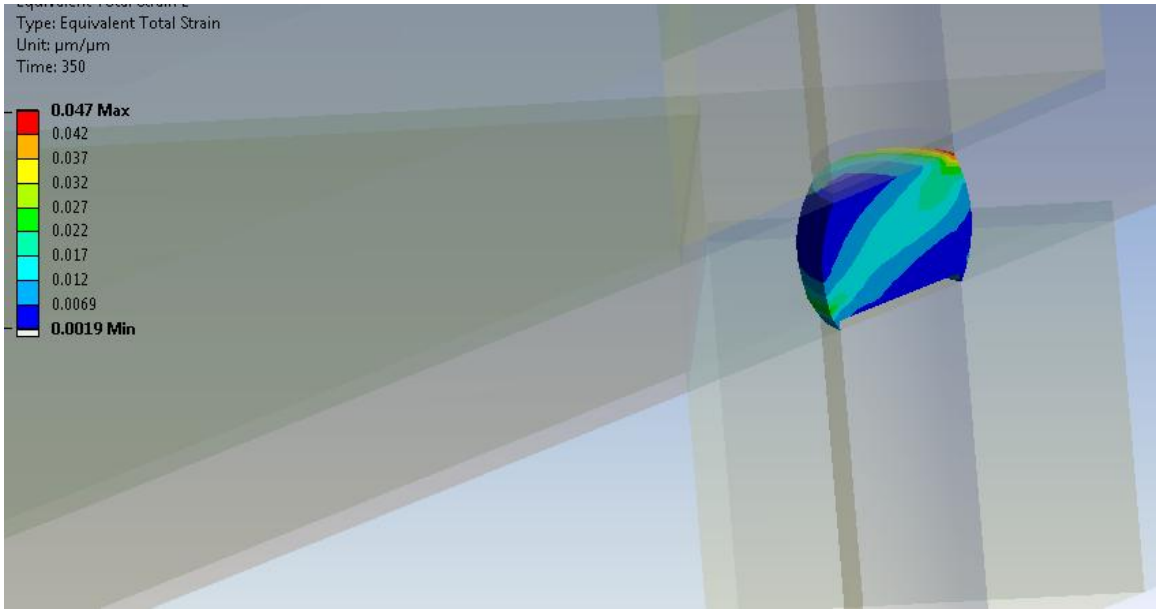


Figure 3-40 Compact model – Equivalent total strain (mold E – 25GPa)

The contour plots (Figure 3-33 through Figure 3-40) demonstrate that the damage in the corner solder joint is in good quantitative and qualitative agreement between the array and the compact model (within 10%). It also shows that the mechanical response of the corner joint is qualitatively in line with the existing literature [61].

Table 3-11 Results Summary – Effect of overmold stiffness

Case	Overmold Elastic Modulus – E (GPa)	Warpage (µm)		Corner joint stress (MPa)		Corner joint strain (%)	
		Array	Compact	Array	Compact	Array	Compact
1	5	665	656	24.5	26.1	2.56	2.4
2	25	1120	1142	32.9	34.4	5.19	4.7
3	40	1303	1307	34	35.3	5.7	4.85

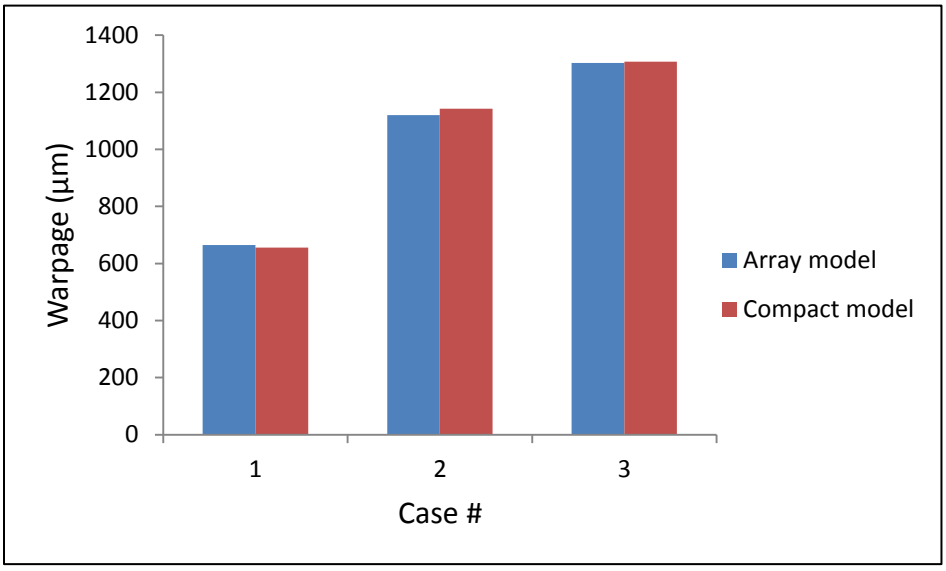


Figure 3-41 Effect of overmold stiffness on warpage

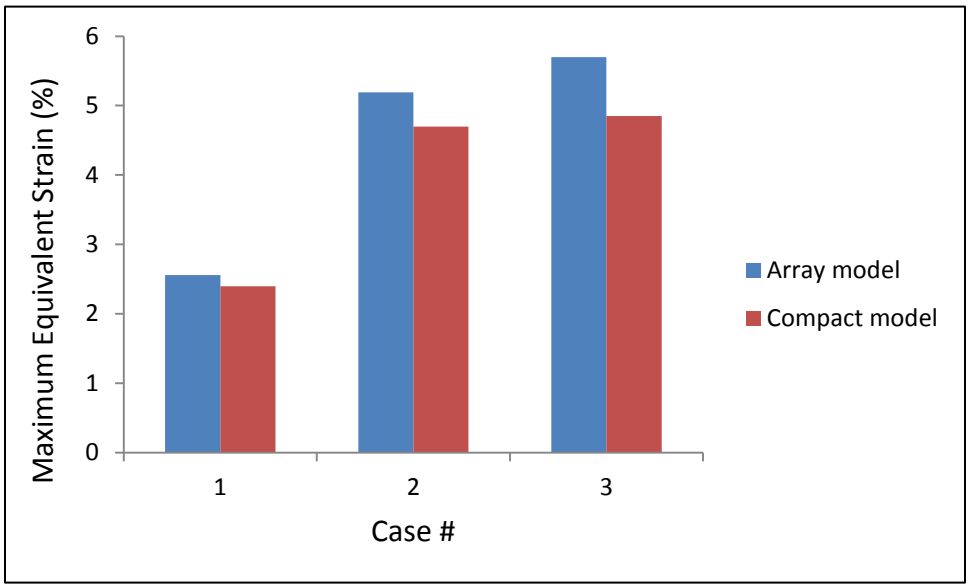


Figure 3-42 Effect of overmold stiffness on corner joint strain

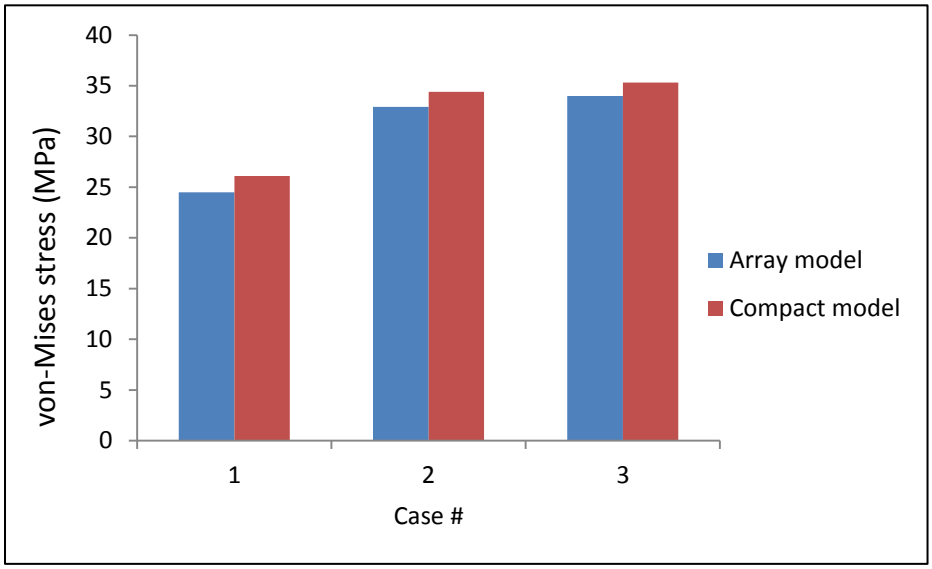


Figure 3-43 Effect of overmold stiffness on corner joint stress

3.7.4 Validation of the compact modeling technique for different loading conditions (Accelerated Thermal Cycling - ATC)

Accelerated thermal cycling is a widely used reliability testing protocol during development, qualification and sustainability. The purpose of the test is to mimic an accelerated stress loading to determine the mechanical integrity of the various package components (especially solder interconnects). CTE mismatch between the different package components (silicon die, substrate and the PCB) causes the assembly to bend thereby inducing significant stresses (normal and shear) on the weak solder interconnections. The response from the accelerated thermal cycling test includes the von-Mises stress and the total strain, which is typically used to compute the plastic work (strain energy) the corner joint has undergone. Eventually, the accumulated plastic work data is used to find the cycles to failure via different fatigue models proposed by Darveaux, Mansion, and Syed [62, 63, 65].

A typical accelerated thermal cycling profile is shown in Figure 3-44. The package was subjected to accelerated thermal cycling from -40C to 125C for 2cycles. Both the array and the compact models were analyzed and compared. The correlation parameters used were the package warpage, stress-strain distribution in the corner joint, and the accumulated plastic work after 2 cycles.

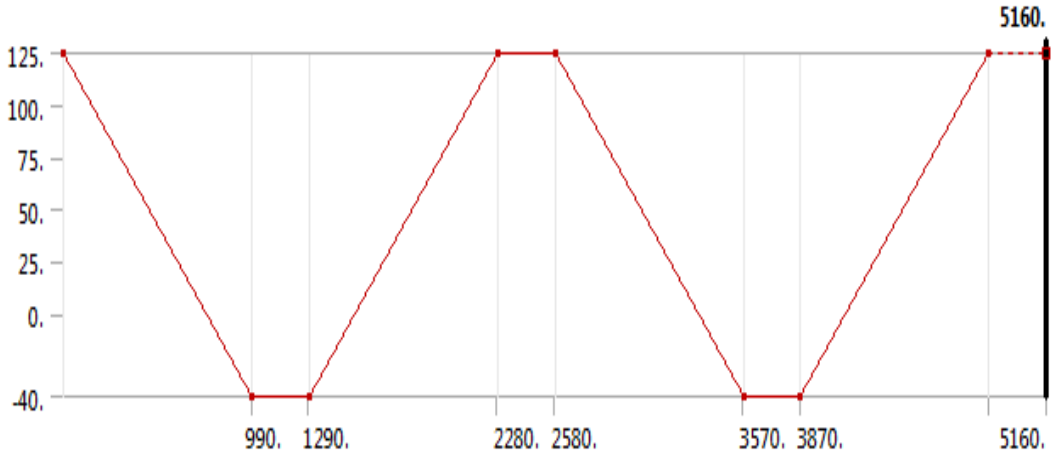


Figure 3-44 Accelerated thermal cycling test profile

For the thermal cycling comparison, case 3 (Table 3-4) was analyzed. Both the array and compact model were subjected to same loading (see Figure 3-44) with 10°C/min ramp rate. The model was considered stress free at 125°C for both the configurations. Same set of effective block material constants were used as shown in Table 3-10

Figure 3-45 through Figure 3-50 show the simulation contour plots for both the array and the compact model. They are in very good qualitative agreement (stress-strain distribution) with a maximum 13% error in the magnitude of the predicted equivalent strain.

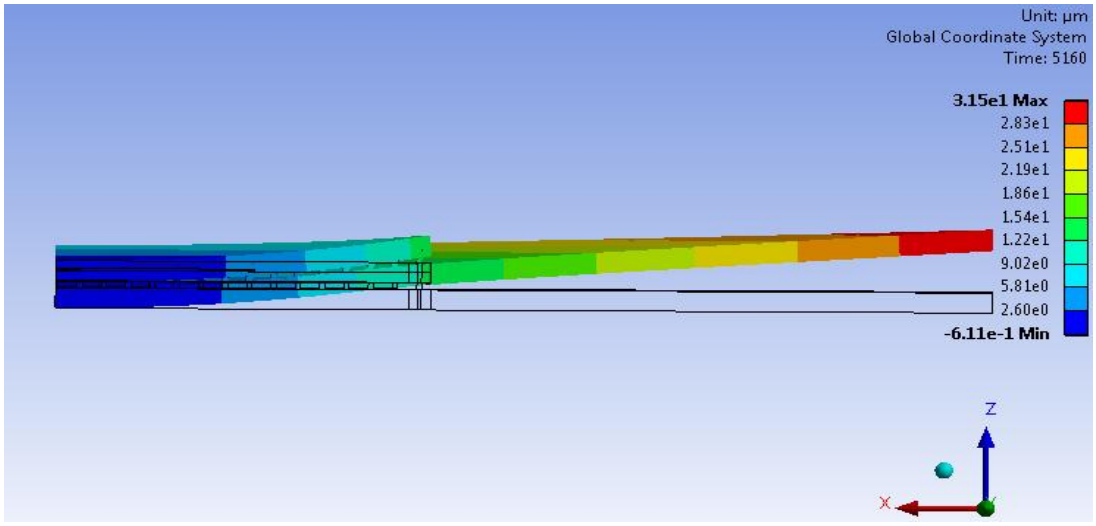


Figure 3-45 Array model Warpage

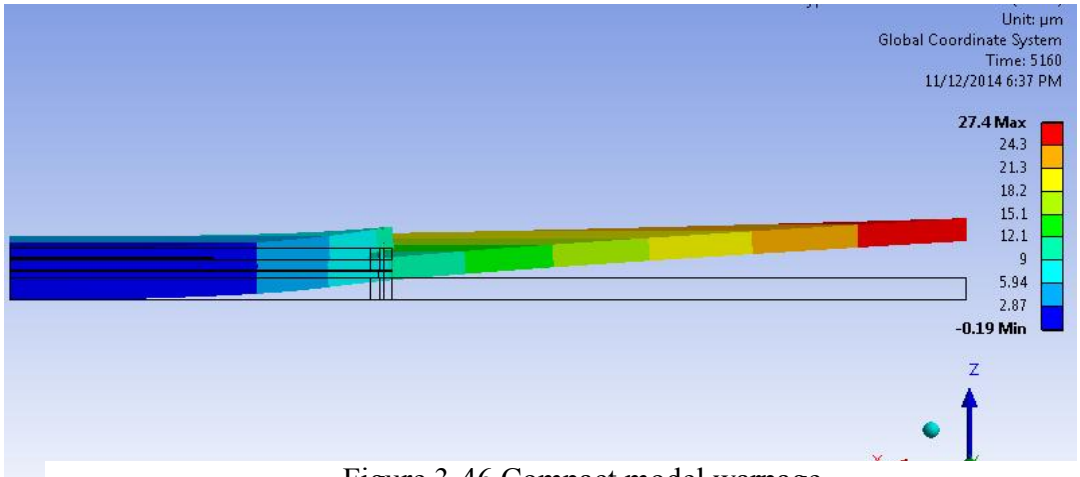


Figure 3-46 Compact model warpage

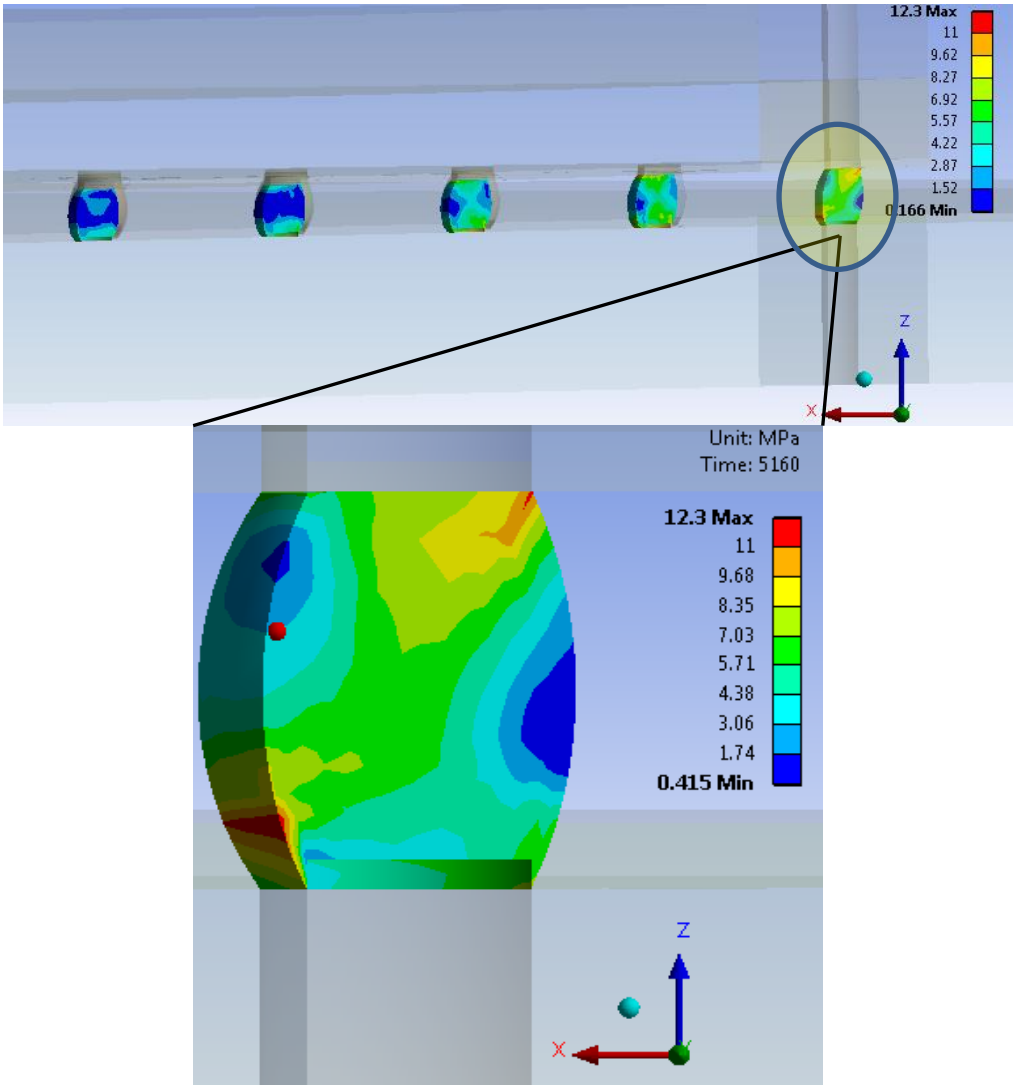


Figure 3-47 Array model – corner joint von Mises stress

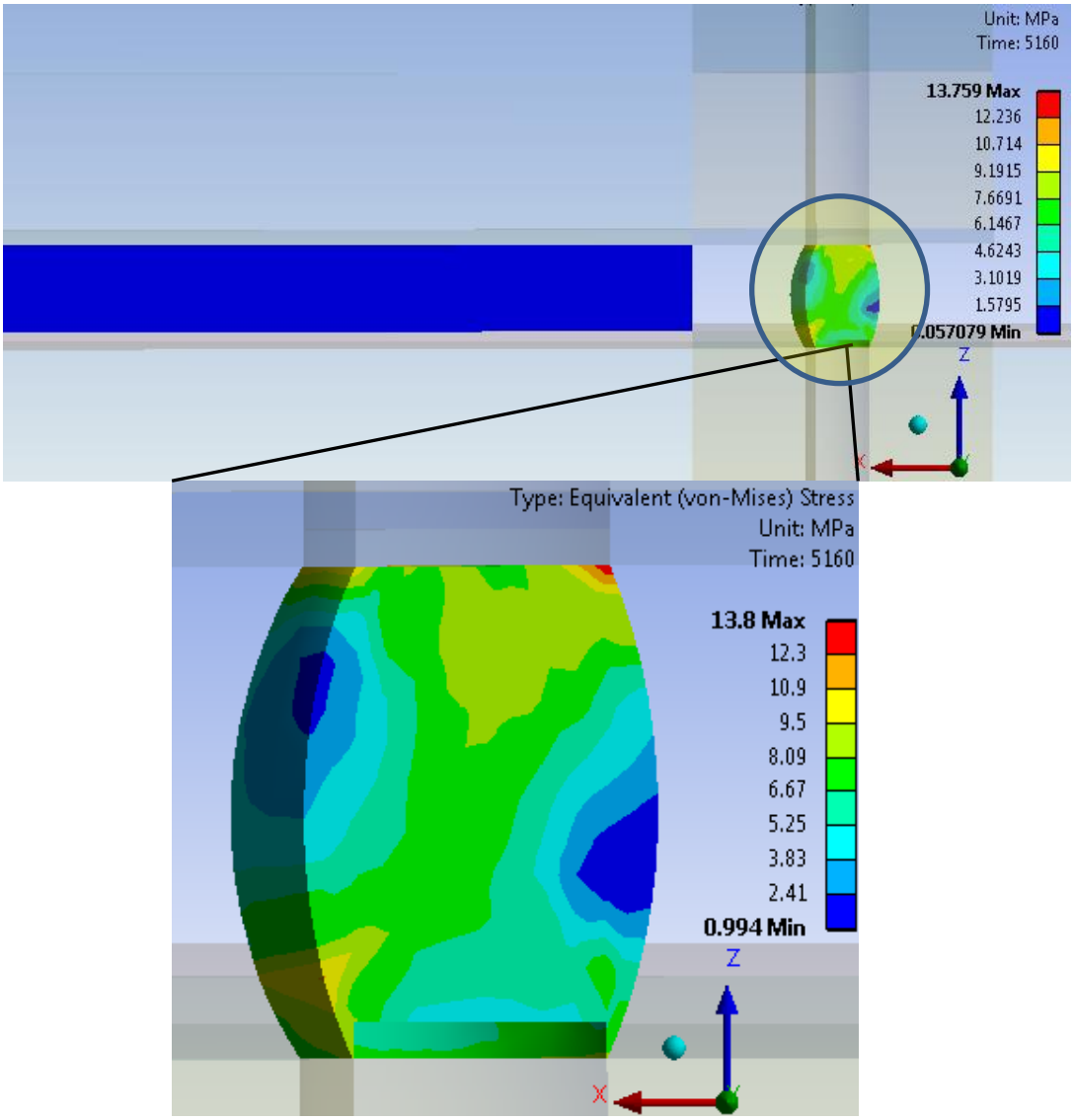


Figure 3-48 Compact model – corner joint von Mises stress

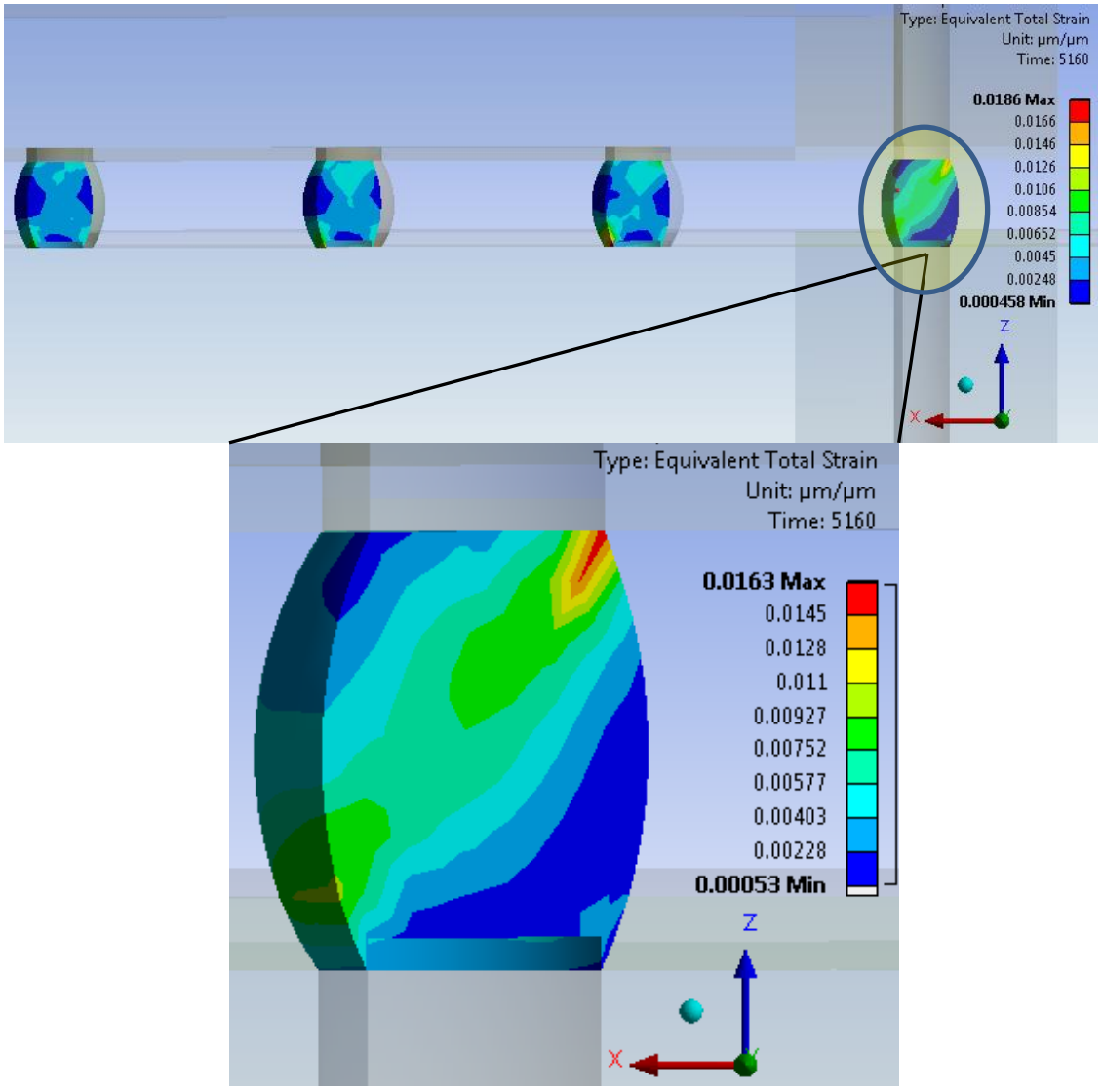


Figure 3-49 Array model – corner joint equivalent strain

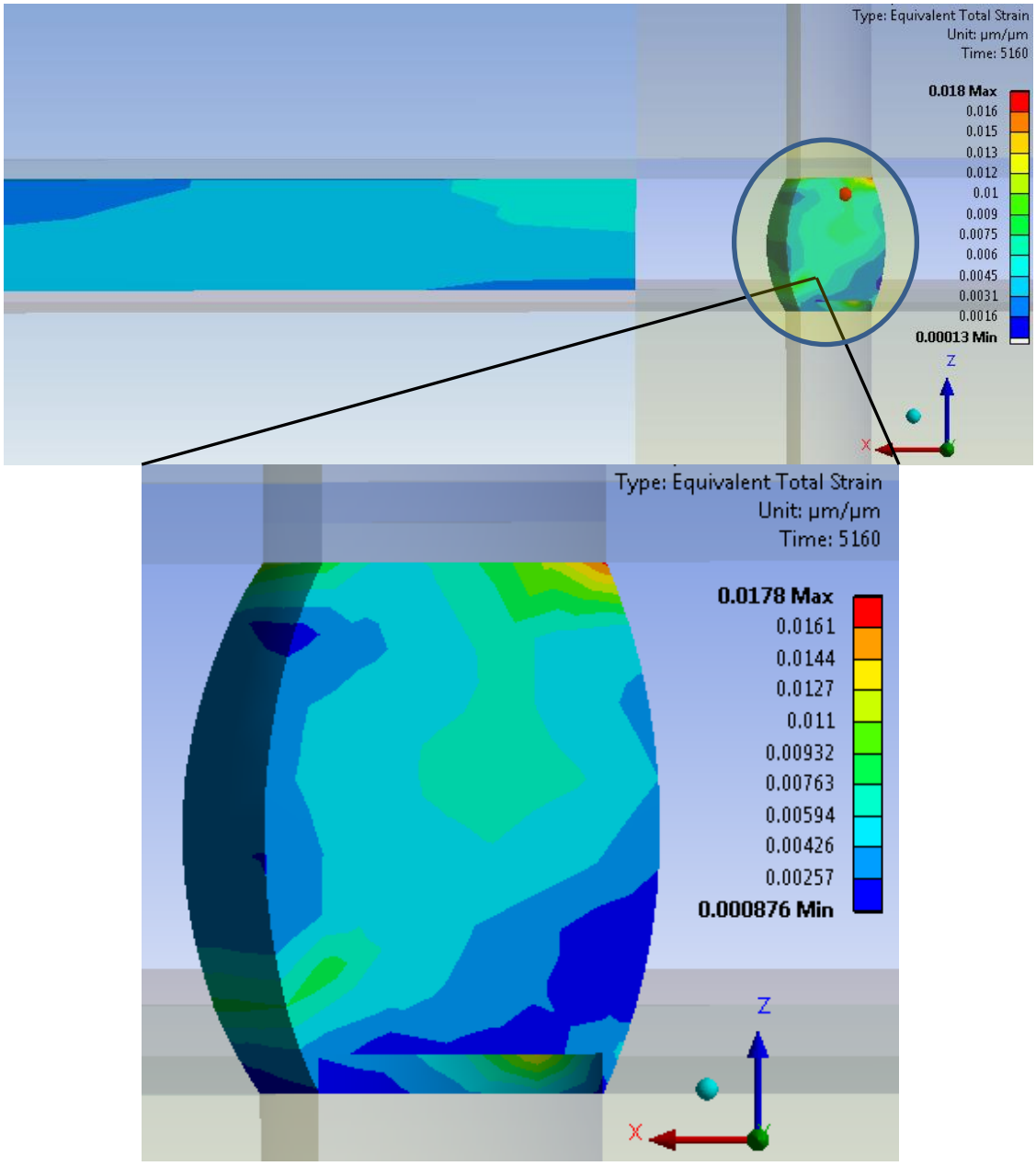


Figure 3-50 Compact model – corner joint equivalent strain

Table 3-12 Accelerated Thermal Cycling Results

Uz – Warpage (μm)			Corner joint stress – (MPa)			Corner joint strain – (% strain)		
Array	Compact	% Error	Array	Compact	% Error	Array	Compact	% Error
31.5	27.4	13	12.3	13.8	12	1.63	1.78	9

Table 3-13 Accelerated Thermal Cycling Results – Damage Correlation

Plastic Work - Damage (MPa)		
ΔW (Cycle 2 - Cycle 1)		
Array Model	Compact model	% Difference
.0565	.0567	9

3.8 Validation of the compact modeling technique across different solder alloys – Sn63Pb37

The compact modeling technique was also validated for Sn63Pb37 solder alloy and it is shown that the proposed technique works across alloys, however, tuning of the ANAND's model needs to be carried out whenever there is a new solder at hand. The contour plots given below show a very good agreement in the damage parameters between the array and the compact model. Typically, SnPb alloy is softer than Pb-free, hence higher strain is seen for the SnPb (qualitative validation of the numerical simulation). It also shows that the compact modeling technique is scalable across different type of solders, however – the constants for the effective block will have to be tuned whenever the solder alloy is changed. A non-dimensional correlation of the new constants with the solder volume fraction will have to be performed in order to formulate a relation similar to that shown for SAC 305 (see Figure 3-31and Figure 3-32).

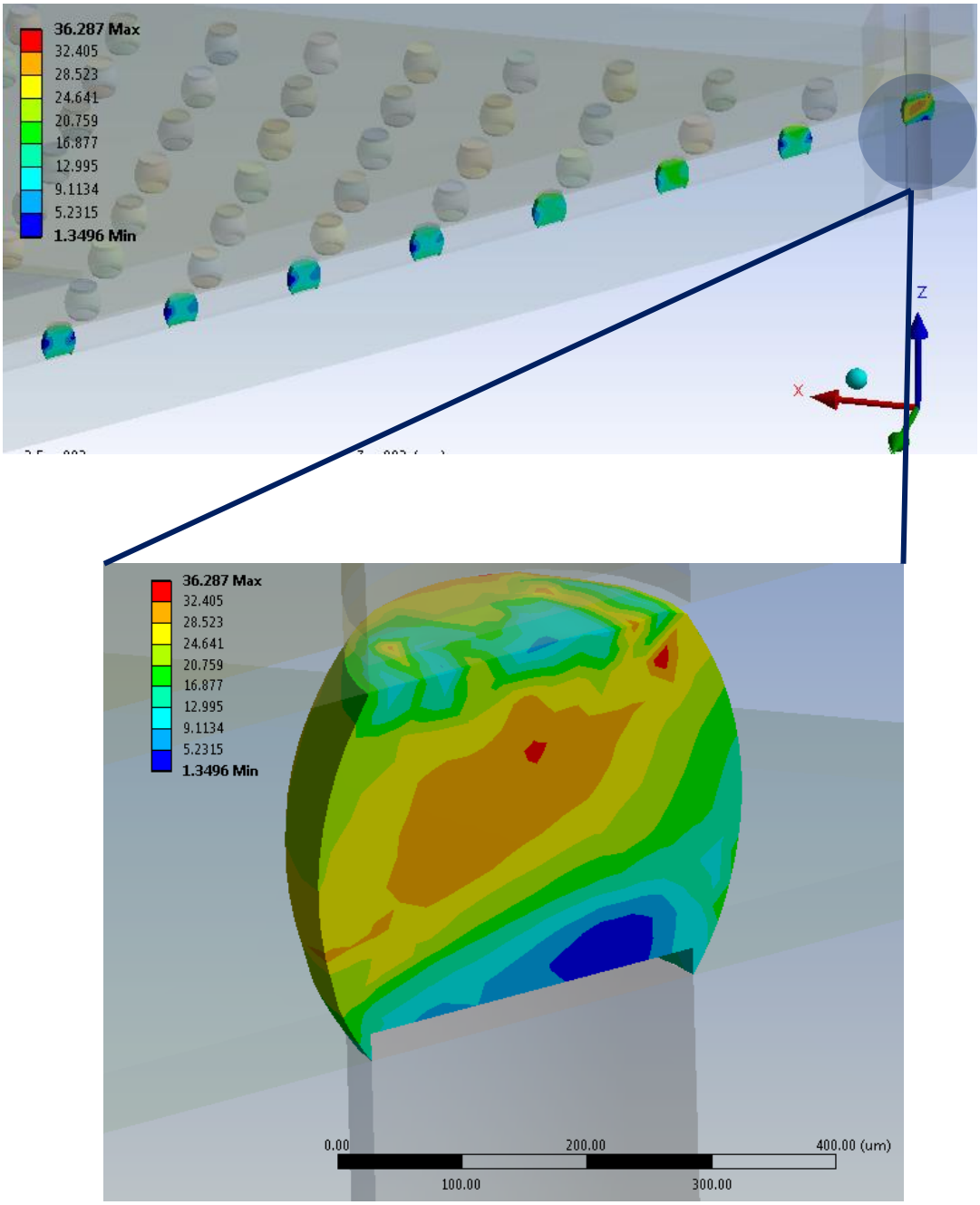


Figure 3-51 Array model – von Mises stress (Sn63Pb37 alloy)

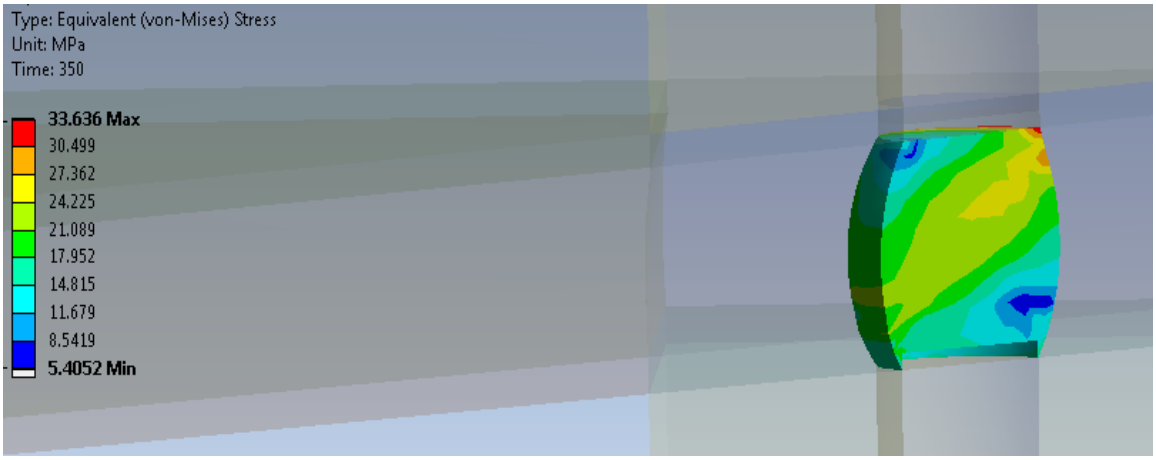


Figure 3-52 Compact model – von Mises stress (Sn63Pb37)

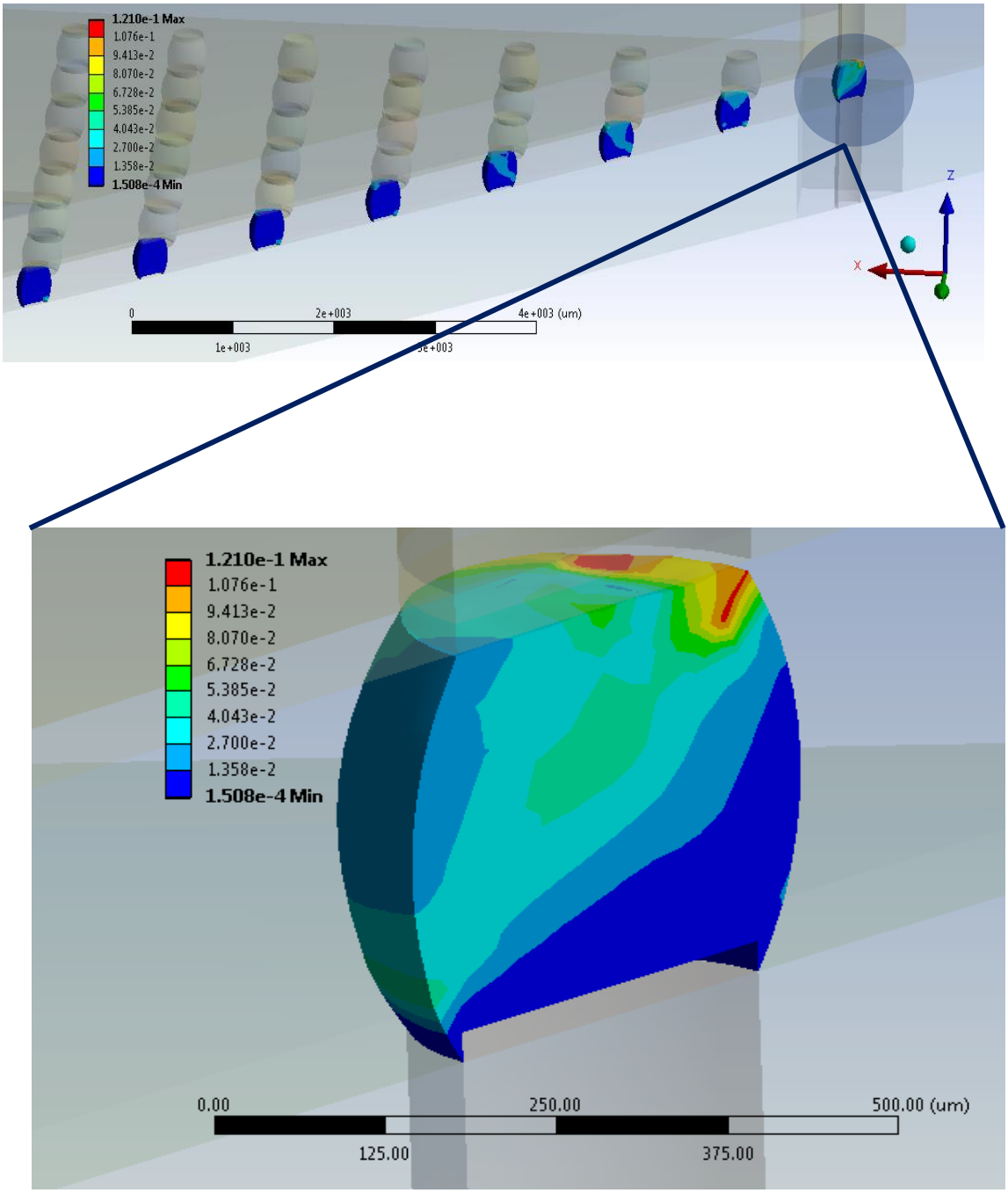


Figure 3-53 Array model – Equivalent total strain (Sn63Pb37)

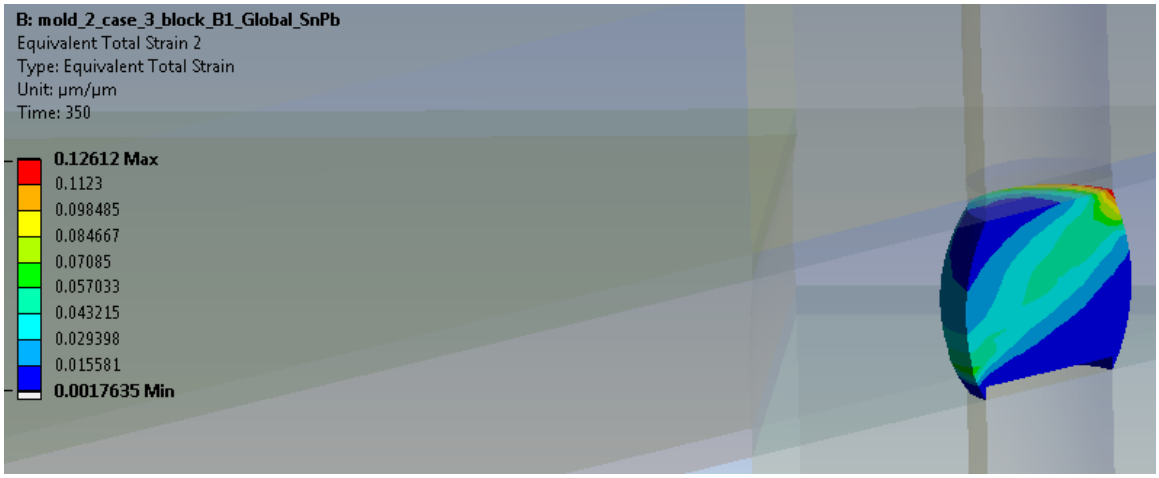


Figure 3-54 Compact model – Equivalent total strain (Sn63Pb37)

Table 3-14 Results summary for Sn63Pb37 solder alloy

Warpage			Corner joint stress (MPa)			Corner joint Strain (MPa)		
Array Model	Compact Model	% Error	Array Model	Compact Model	% Error	Array Model	Compact Model	% Error
1221	1294	6	33.6	36.3	7.5	12.1	12.6	4.1

Chapter 4

3D TSV PACKAGE - COMPACT MODELING METHODOLOGY DEVELOPMENT FOR CPI ANALYSIS

4.1 3-D: Merits and Challenges

The convergence and miniaturization of computing and communications dictates building up rather than out. As planar device miniaturization continues to its ultimate limits, the complexity of circuit interconnections for 2-D devices becomes a limitation for performance and drives up power dissipation [1]. As the consumers demand more functions on their hand-held electronic devices, the need for more devices such as memory, CPU and GPU in hand-held type footprints is increasing. This results in high package density. Chip-stacking (3-D) is emerging as a powerful tool that satiates such IC package requirements. Access to the third dimension has significantly simplified chip-level communications and transfer of information among the processing elements and has provided rapid access to memory and configurable logic. 3-D technology would enable extremely dense solid-state memory to be arrayed within a few microns of the processing elements, which reduces access times. The 3-D arrangement also provides opportunities for new circuit architectures based on the geometrical ability to have greater numbers of interconnections among multi-layer active circuits. A 3-D FPGA would overcome the interconnect limitations, resulting in greater silicon efficiency per function (number of used gates/total number of gates), faster signal/data throughput, and faster switching of the gate-level configuration. True 3-D ICs can operate at higher clock rates and can

consume less power over their 2-D implementations, as the 3-D arrangement minimizes the length of circuit interconnects (see Figure 4-1) [35].

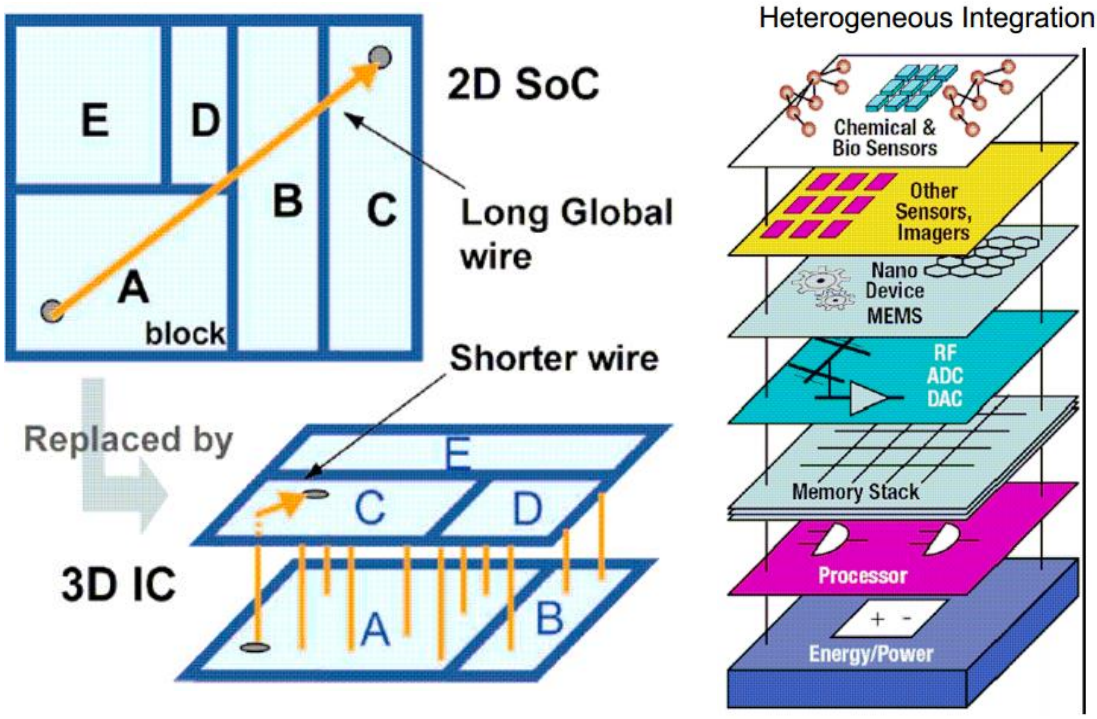


Figure 4-1 2D vs 3D

The obvious electrical/performance benefits make 3D IC technology very attractive for future semiconductor products, especially handheld. However, there are some research and implementation challenges that are not yet fully understood. Since die stacking leads to increased power density within the same footprint, it is clear that heat removal from a multi-die stack is going to be far more challenging than that of the traditional ICs [36, 18]. Integration in the Z direction is achieved by stacking dies and interconnecting them with WB (wire bonding), FC (flip chip) or TSV (through- silicon-

via) [19]. TSV is one of the key enabling technologies for 3-D systems. 3-D TSV technology is being termed as the "***next big thing***" in the semiconductor arena and has the potential of revolutionizing the packaging industry but it has some inherent issues that need to be addressed before it could be implemented in the mainstream electronics industry. TSV fabrication process, thermal management of 3-D TSV packages, TSV joule heating, and chip package interaction (CPI), are some of the key issues in this technology [37, 38, 39, 40, 41]. In this chapter, the thermo-mechanical CPI analysis is carried out and a full field compact 3D modeling methodology has been proposed. This modeling methodology will result in stress predictions at the Si/TSV region and inter die μ-bumps during attachment to the substrate.

4.2 Introduction/ Literature Review

3-D stacking of the processor and memory components in high computing applications reduces the communication delay in a multi-core system owing to reduced system size and shorter interconnects [18]. Stacking 24 devices using special etching technology that connects the layers electronically by punching holes from the highest layer to the bottom layer were developed by *Samsung* in 2013 resulting in much higher density 3-D V NAND flash memory. Samsung's 3D V-NAND is able to provide over twice the scaling of 20nm-class planar NAND flash [16]. High-density-interconnects (HDI), necessitates area array packaging with much higher number of interconnects as well as reduced footprints leading to 3D TSV technology. *Micron's* hybrid-memory-cube is one such development [17]. HMC is a small, high-speed logic layer that sits below

vertical stacks of DRAM die that are connected using through-silicon-via (TSV) interconnects.

TSVs allow 3-D chips to be interconnected directly and provide high speed signal processing. 3-D IC technology is a promising approach to reduce interconnection power, increase communication frequency, improve design flexibility and enable seamless system-level integration of heterogeneous technologies. In 3-D technology, chips are stacked vertically in the Z-direction providing a volumetric packaging solution. Due to its reduced size, reduced device cost resulting from increased yield opting SOC devices into several smaller area chips, higher electrical performance, and more design freedom for fabricating novel form factors, 3D is gaining popularity in the electronics industry. The GLOBALFOUNDRIES 3D innovation roadmap for high- performance computing is shown below (see Figure 4-2) [42].

Zheng *et al.* [19] introduced the development trends of 3-D stacked packages. The advantage of 3-D over the traditional 2-D or planar packaging (Multi-Chip-Module, MCM) and challenges faced by the 3-D technology have been discussed. Consumer electronic products such as digital cameras, personal digital assistants, cell-phones etc., require high functional integration in small footprints with low cost. Multi-chip packaging (chips packaged on the same plane) is one of the solutions. But, due to miniaturization, coupled with the requirements of high memory density, performance, and more features per cm^2 of Printed Circuit Board (PCB), engineers have been forced to think vertically. Stacking dies and interconnecting them vertically accomplishes all of these goals.

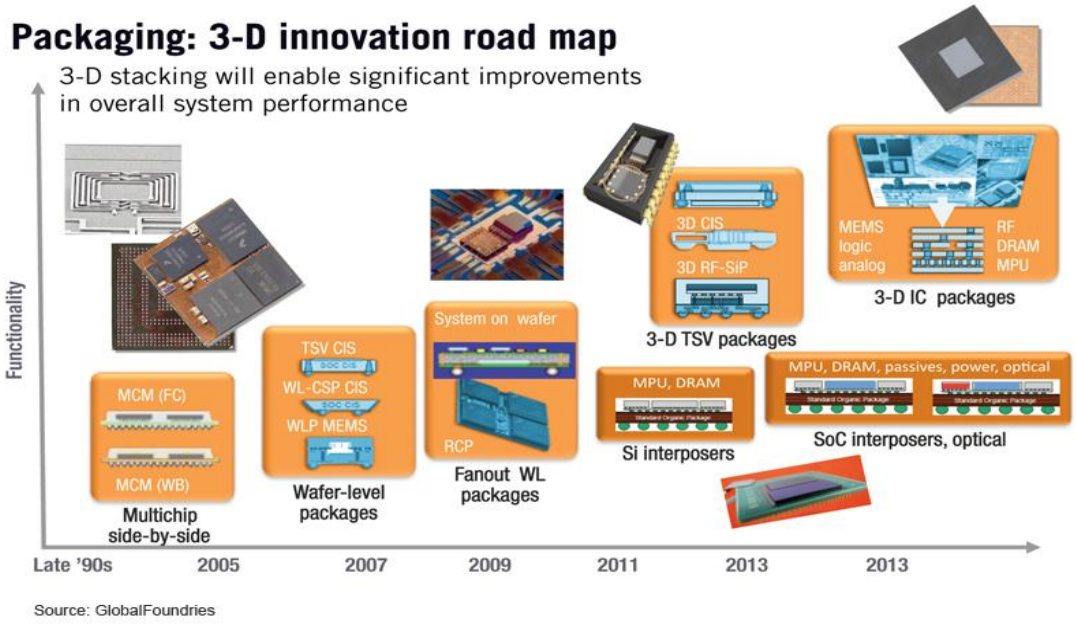


Figure 4-2 3D innovation roadmap - GLOBALFOUNDRIES

Selvanayagam et al. [37] demonstrated the nonlinear stresses and strains in the μ-bumps between the silicon chip and copper filled TSV interposer (with and without underfill) for a wide-range of via sizes and pitches, and various temperature conditions. Selvanayagam [37] results were useful for making a decision if underfill is necessary for the reliability of μ-bumps and selecting underfill materials to minimize the stresses and strains in the μ-bumps. [38] et al. developed an analytical model for the three-dimensional state of stress in a periodic array of TSVs. The model accounts for Cu plasticity and predicts the out-of-plane protrusion that occurs in the TSV due to differential thermal expansion with the surrounding Si. Excessive out-of-plane deformation of the top surface of the via has the potential to induce fracture causing stress in the brittle dielectric layers that lie above the via. Kawa et al. [39] studied the

TSV related performance and reliability issues due to thermal-mechanical stress using Fammos TX. He showed that the localized TSV/Si stress significantly affects the BEoL mechanical response/reliability and the transistor performance (see Figure 4-3 through Figure 4-5) [39].Figure 4-5 TSV Cu protrusion [39]

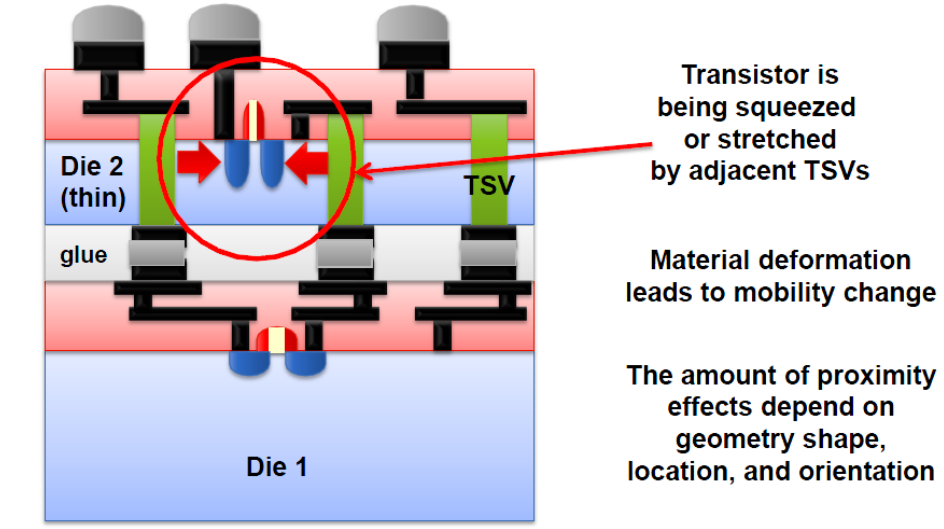


Figure 4-3 Stress affects transistor performance [39]

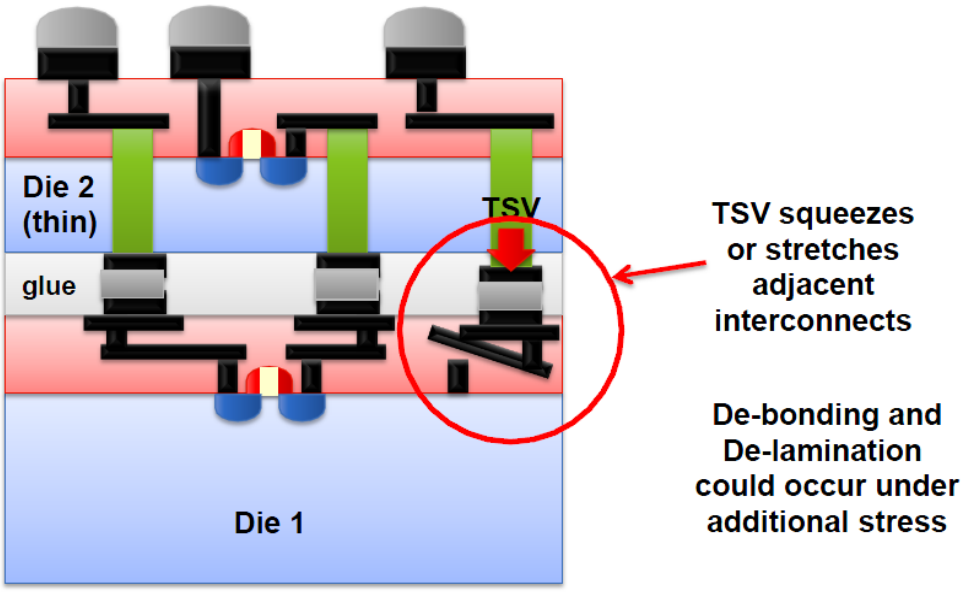


Figure 4-4 Localized stress affects the BEoL stack [39]

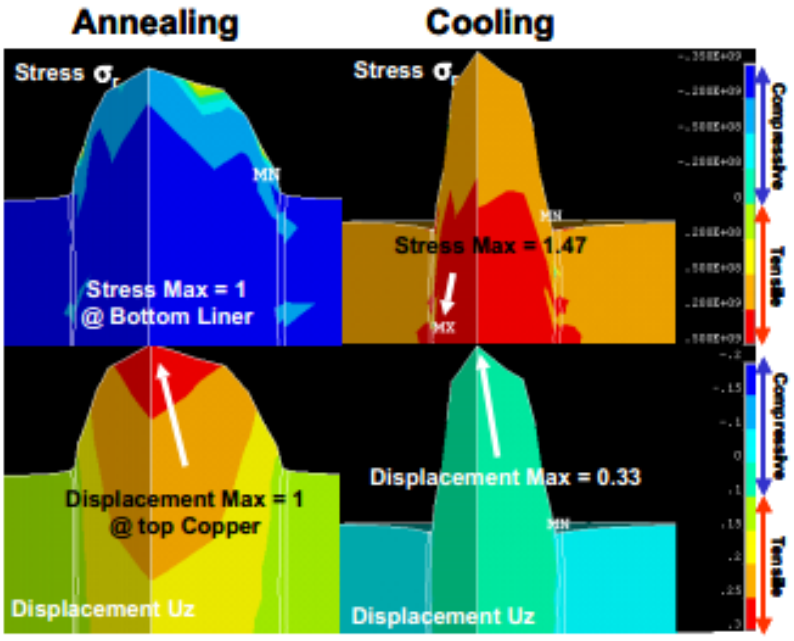


Figure 4-5 TSV Cu protrusion [39]

Alam *et al*. [5] analyzed the parasitic characteristics of inter-die bonding techniques and materials.

Comparing 3D technology with CSP (chip scale package), which has a silicon efficiency of about 80%, 3-D ICs silicon efficiency is greater than 100% without increasing the thickness or the footprint of the package. Integration in the z direction is achieved by stacking dies or stacking packages and interconnecting them with WB (wire bonding), FC (flip chip) or TSV (through silicon via) [19]. For a conventional single chip package, heat can be dissipated through the top (spreader and eventually through the heat sink), bottom (through the substrate and eventually PCB). But for a 3-D package used in mobile products, heat dissipation is even more difficult, as heat sink implementation is difficult due to space constraints. Sung *et al.* [43] studied the thermal distribution in a 3D TSV package with various die stacking architectures. They also analyzed the effect of location and number of TSVs on the maximum temperature of the die. Maximum temperature decreased with TSVs but it was independent of via location.

A lot of work is being done in the area of cooling of 3-D chip stacks. Matsumoto *et al.* [44] discussed and evaluated various possible cooling methods from the bottom and periphery of the silicon interposer. Kota *et al.* [45] presented a parametric study focusing on the design and thermal properties of a liquid interface thermal management solution for 3-D stacks using a radial heat sink cooled by an array of synthetic jet actuators. Phan *et al.* [46, 47] developed a novel cooling method for 3-D ICs using a multi-dimensional thermoelectric cooler. Moon et al. [40] presented a thermal management of a stacked-die package in a handheld electronic device using passive solutions.

As stated earlier, device performance is a function of several parameters and varies with the junction temperature and the performances decreases with increase in temperature. Karajgikar *et al.* [48] showed correlation between the temperature and performance for a 90nm technology Pentium IV Northwood architecture. Electron mobility (e^- mobility) variation with doping concentration and temperature is shown below [49] (see Figure 4-6). Electron mobility is a measure of how fast the electrons would move when excited by a voltage drop, i.e. e^- drift velocity. Mirza et al. [50] demonstrated the effect of junction temperature and TSV diameter on the interconnect delay.

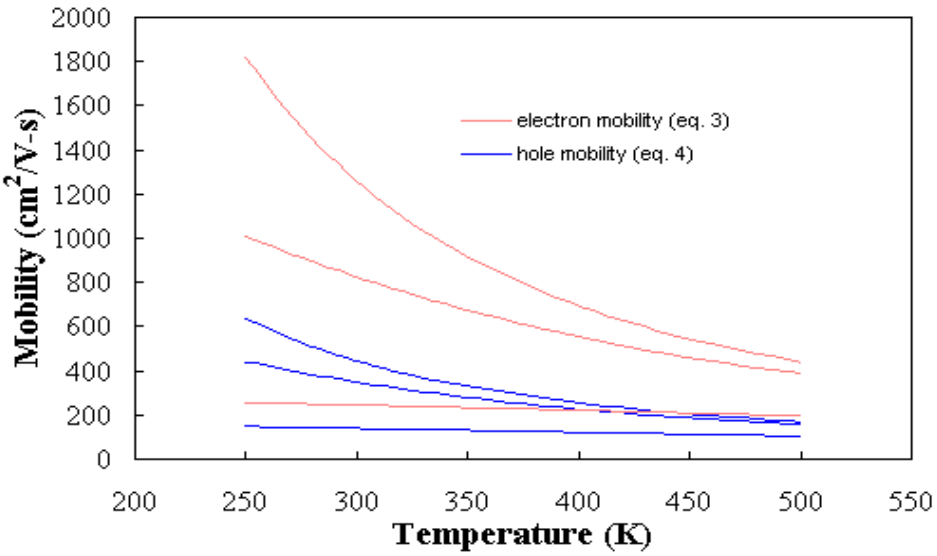


Figure 4-6 Mobility variation with temperature/doping concentration

There is a need to determine if TSV joule heating (caused when current flows through it) is significant enough to cause a substantial change in the junction temperature and eventually lead to performance degradation. Based on an NSF project, the effect of

TSV joule heating on the device performance measured by trans-conductance, electron mobility (e^- mobility), and channel thermal noise is studied for thinned (100 μm) dies with a *uniform* power map. Results indicate that joule heating has a significant effect on the thermal response of the 3D IC and subsequently proves to be quite detrimental to the over-all performance.

Full field CPI analysis of a 2 die 3-D flip chip package with TSVs is performed to study the thermo-mechanical response during chip attachment to the substrate. TSVs occupy approximately 1.5% of the chip real estate ***(CRE)*** and the TSV diameter is 10μm including the 1μm thick dielectric around it. We limited the TSV area to <4% so we don't inadvertently affect the silicon efficiency. Since there is a huge scale difference among the different package components and the full model size is big (~8000 TSVs and μ-bumps) a novel compact modeling technique has been implemented to maintain reasonable computational time. Simulation is performed in 2 steps – a compact global model is formulated and solved; results from the compact global model are used as boundary conditions in the sub-model which is usually the far corner region of the package (critical region) with detailed features such as actual μ-bump interconnects and the TSVs. The proposed compact model has the corner region modeled in detail and replacing the inner region of Si/TSVs with an effective Si block (with volume averaged material properties) [34]. A similar effective block for the 1st level interconnects (μ-bumps) has been modeled for the inner region while the corner μ-bump is modeled in detail. The concept of effective block is similar to that explained in chapters 2 and 3 earlier. Thermo-mechanical response of the inter-die bond layer (μ-bumps) is analyzed

and the “effective block” model is compared and validated with the “full array” version for a smaller footprint.

4.3 Model Description

A 3D TSV package is modeled to demonstrate the feasibility of the compact modeling technique for CPI study of the mechanical integrity of *TSV/SiO2* and μ-bump region during package attachment to the substrate. The package footprint is based on a HDI mobile application package (memory-on-logic) footprint (see Figure 4-7) [57]

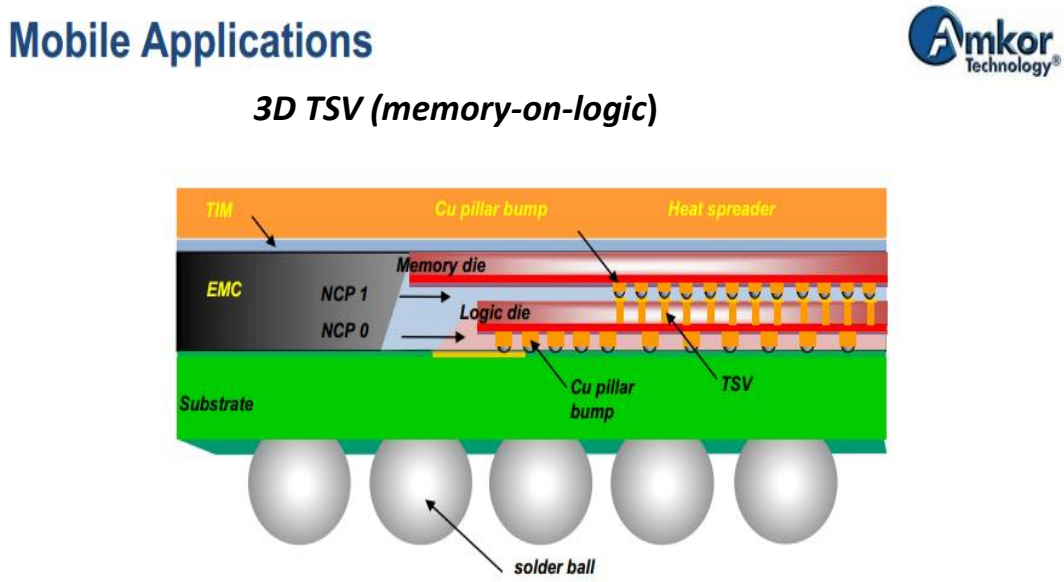


Figure 4-7 Package footprint [57]

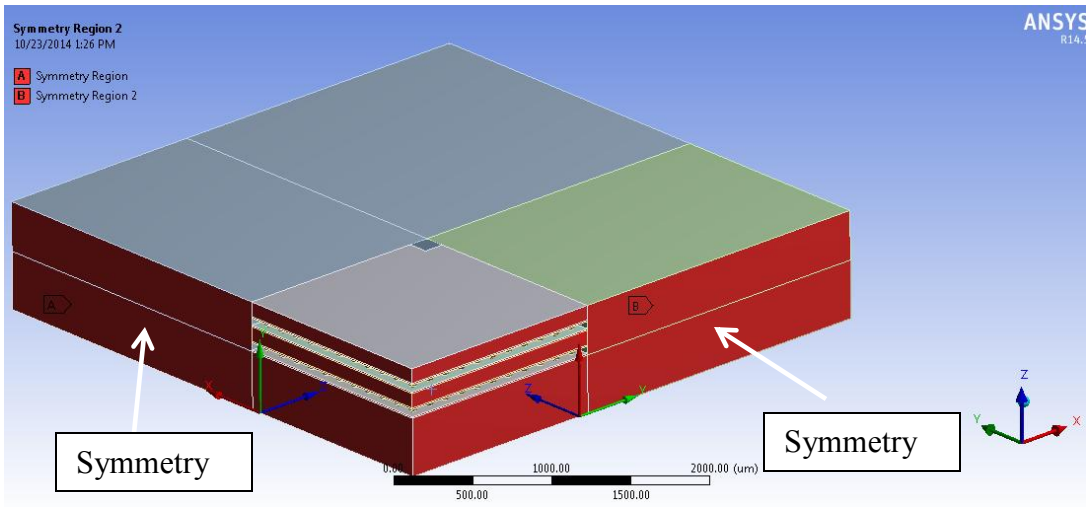


Figure 4-8 Quarter symmetry model – symmetry faces

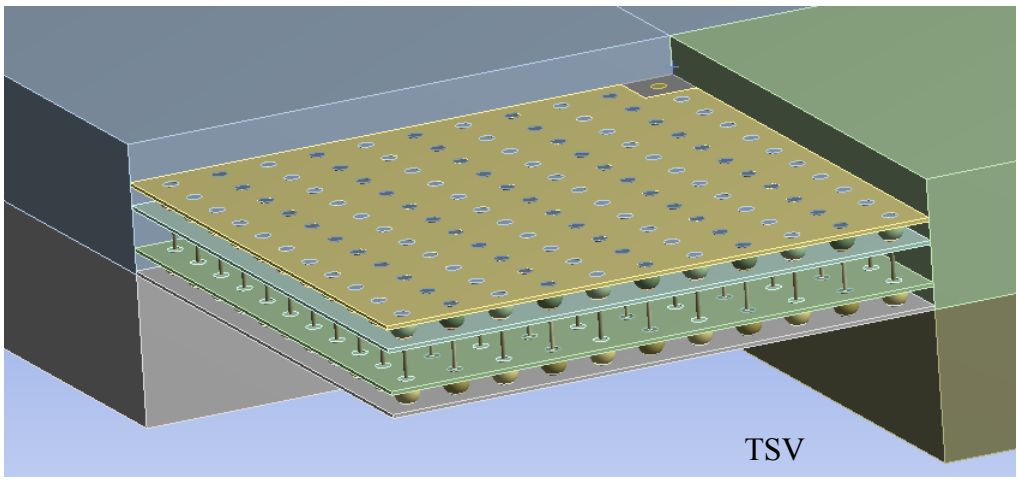


Figure 4-9 Quarter symmetry model – internal view

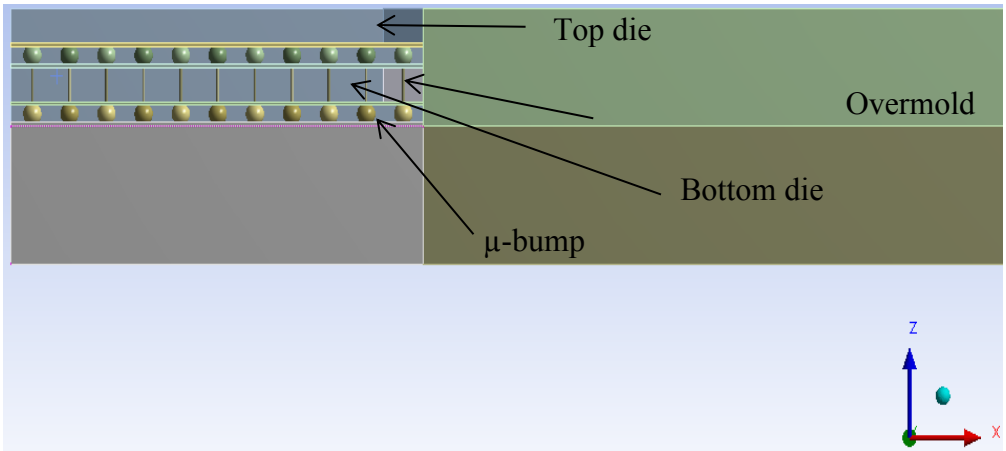


Figure 4-10 Package Cross-section

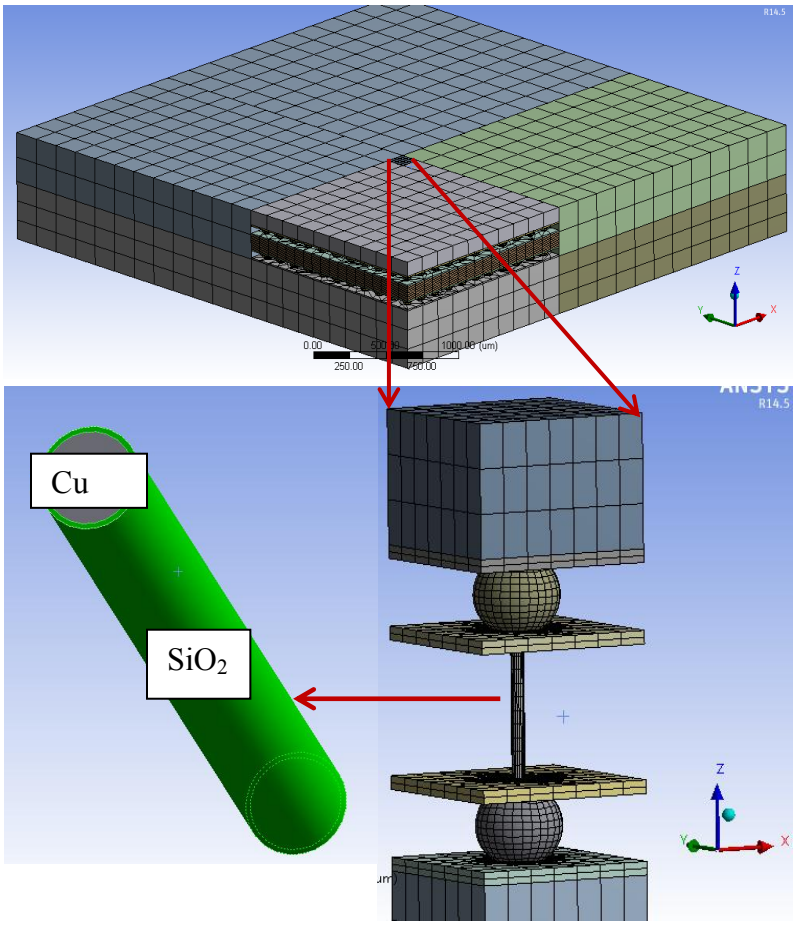


Figure 4-11 Corner region detailed view

Figure 4-8 through Figure 4-11 show the ANSYS models for a smaller array footprint (for practical computational times). As shown in Figure 4-12 sub-modeling has been leveraged to assess the mechanical integrity of the Si/TSV region. In the global model – the TSVs are models as one volume (10um diameter). In the subsequent sub-model, the SiO_2 dielectric (1µm thick) is modeled to capture the interfacial interaction of the Cu/SiO2 region and the Cu region is 9µm in diameter. TSV length is 100µm maintain an aspect ratio of 10.

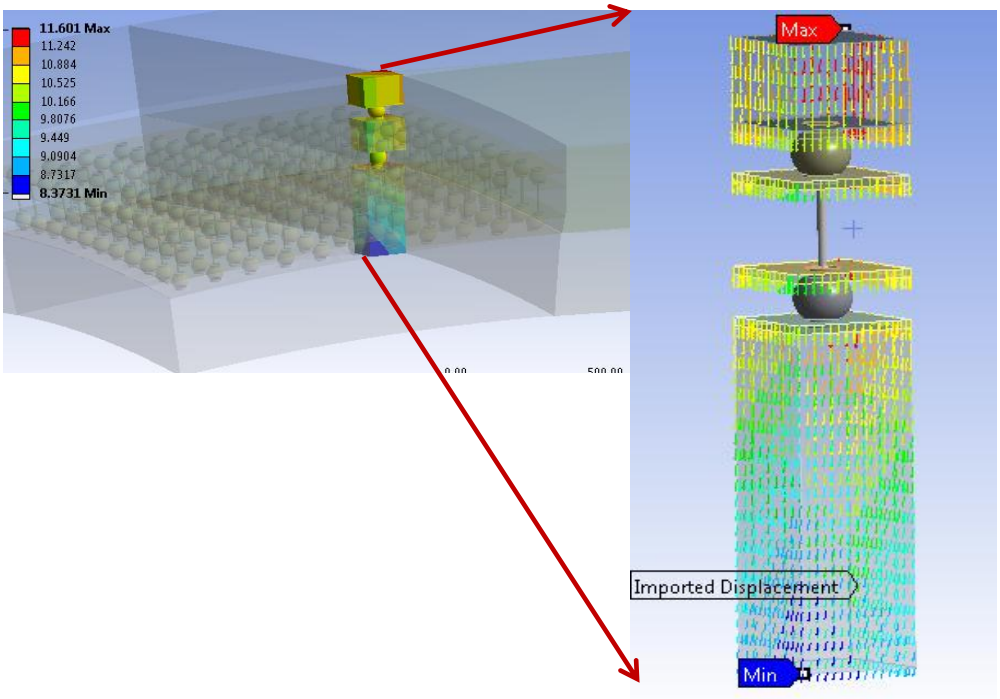


Figure 4-12 Sub-model (corner region)

4.3.1 Material Properties and Dimensions

All the material properties used in this analysis are same as that used in chapters 2 and 3, except for the Cu TSV, SiO_2 dielectric around the TSV and the polyimide. Copper TSV is modeled elastic-plastic as shown in Table 4-1.

Table 4-1 Material Properties

Material	E - GPa	ν	CTE – ppm/°C	Plastic curve for Cu	
				Strain	Stress (MPa)
Cu TSV	121	0.3	17.3	.001 .004 .01 .02 .04	121 186 217 234 248
SiO_2	71.4	.16	.5		
Polyimide	1.2	.34	52		

Table 4-2 Package Dimensions

Component	Thickness (μm)	Length (mm)	Width (mm)	Diameter (μm)	Height (μm)
Silicon Die	100	3	3		
TSV	10	-	-	10	100
SiO2	0.5				
Substrate	400	5	5	-	-
Solder Bumps	-	-	-	80	50
Overmold	400	5	5	-	-

4.3.2 Compact Model Development

A typical mobile application HDI 3-D TSV package consists of thousands of TSVs and µ-bumps with fine pitches. Performing a thermo-mechanical simulation to estimate the warpage and the localized stresses in the TSV region and the µ-bumps is quite cumbersome (computationally expensive) and sometimes impossible. Also, if an isolated study of the TSV/µ-bump region is performed (just 1 unit cell) then the package warpage and the length effect on the damage cannot be estimated. Therefore, there is a need to develop a reasonably accurate modeling methodology that can estimate the mechanical response in the vulnerable region. The proposed methodology showed huge computational time savings with negligible error when compared to the full array model.

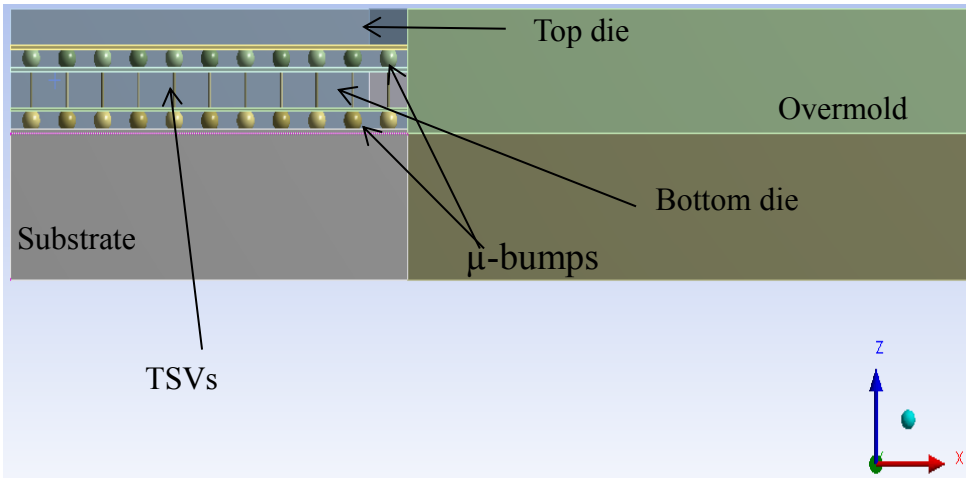


Figure 4-13 Array Model cross-section

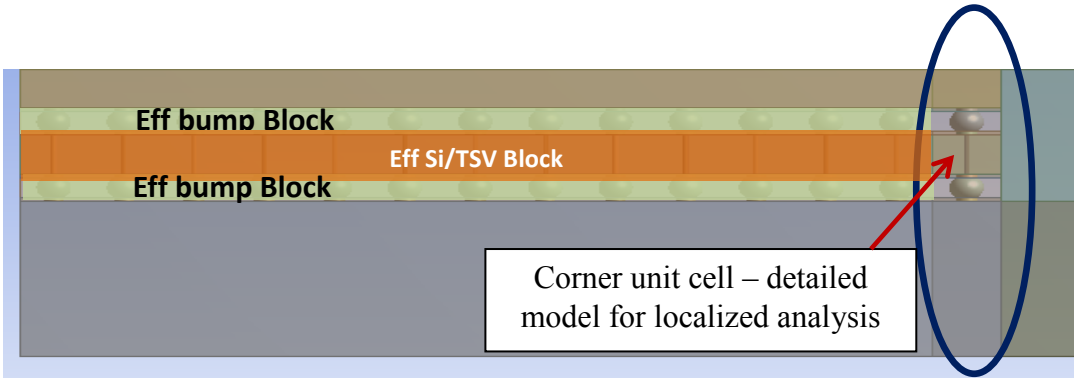


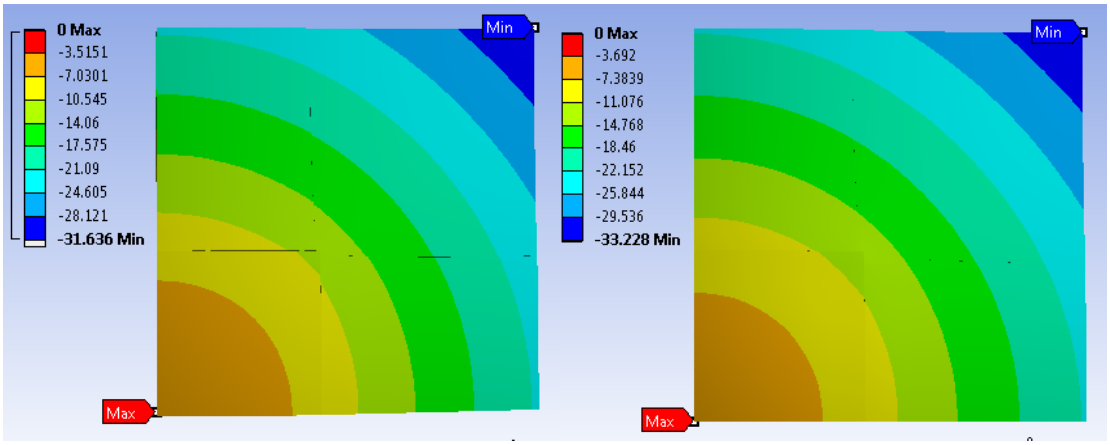
Figure 4-14 Compact model cross-section schematic
(TSVs/µbumps replaced with an effective block)

As shown in Figure 4-14, the µ-bumps and the TSVs are replaced with an effective block with effective material properties (properties determined as demonstrated in chapters 2 and 3). The challenge in developing this modeling methodology is to come up with a material model for the effective blocks (Si/TSV block and µ-bump block) such that the global effect of the respective effective blocks is similar to that of the array in the full array model. The compact model resulted in significant reduction of the mesh size from 140000 elements to 5000 elements (most of the elements reduced were for the viscoplastic solder and elastic-plastic Cu TSV), thereby leading to a huge computational time reduction. To verify the accuracy of the compact model it was compared against different response parameters – maximum package warpage, TSV stresses and strains, µ-bump stress-strain distribution. The comparative results are discussed in the next section given below.

4.4 Results and Discussion

To study the feasibility of the proposed compact modeling technique, the compact model was rigorously validated against its full array counterpart for different output parameters including global package warpage, stress-strain distribution in the μ-bumps (inter die strata), and the stress-strain distribution in the Cu/TSV region. The loading was the reflow condition for the 3D TSV package when it is being attached to the substrate, 200°C to Room (for Pb-free SAC 305 Alloy). Section 4.4.1 talks about the comparative simulation results (ANSYS contour plots) between the "array" and "compact" model demonstrating the feasibility of the proposed technique.

4.4.1 Validation of Compact Modeling with Full Array model



Array Model Warpage Compact Model Warpage

Figure 4-15 Warpage Comparison (top view)

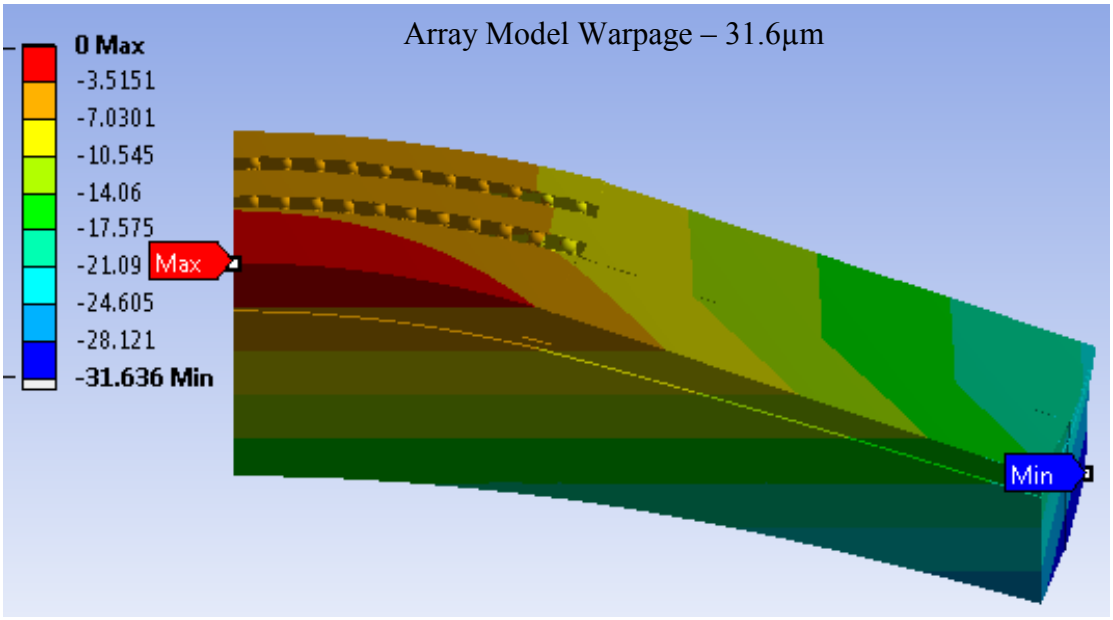


Figure 4-16 Warpage comparison (front view)

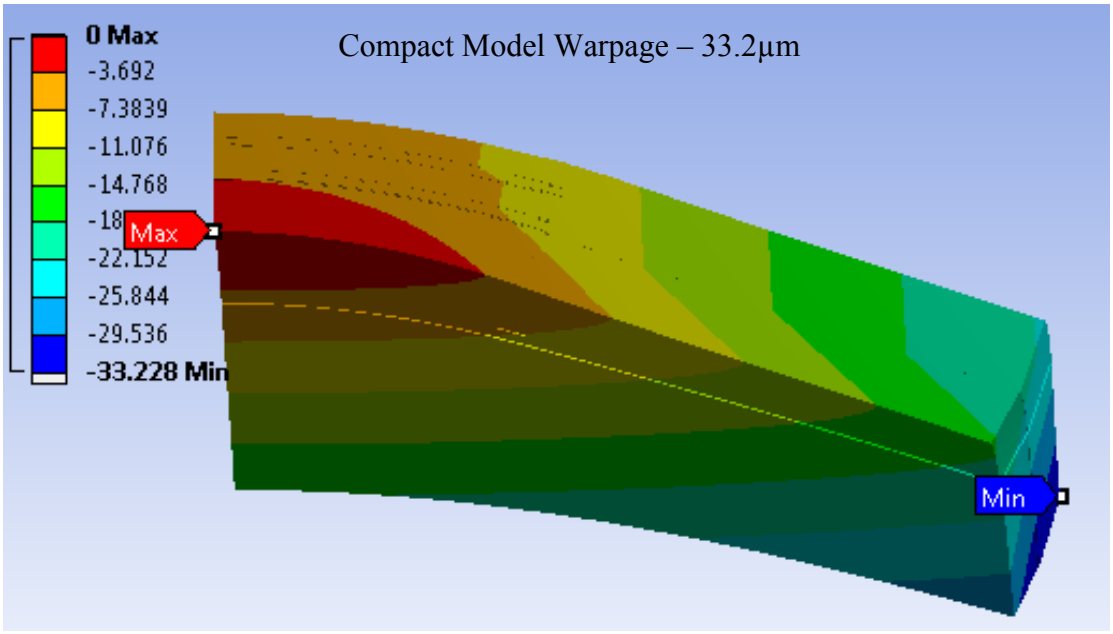


Figure 4-17 Warpage – compact model (front view)

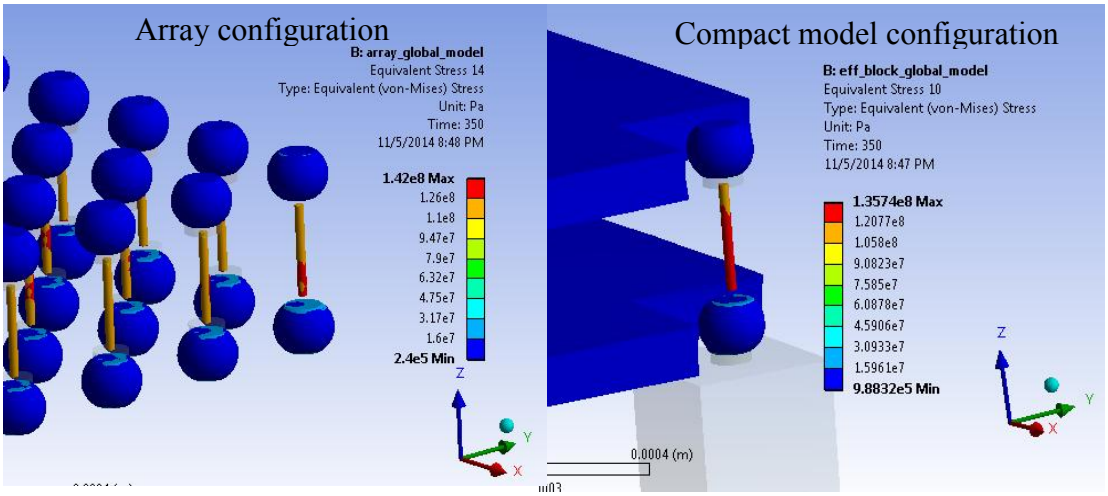


Figure 4-18 Global model TSV von mises stress distribution

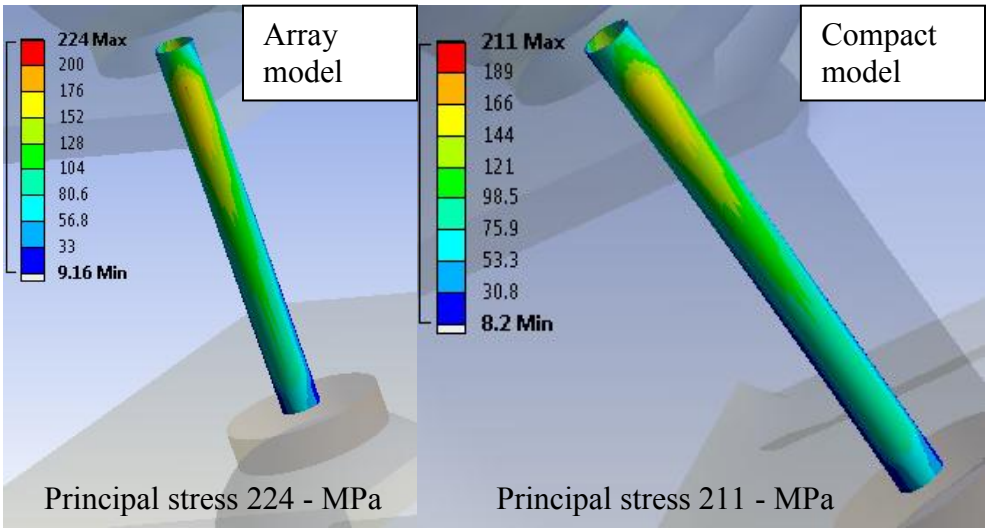


Figure 4-19 Principal stress – SiO_2 region (sub-model)

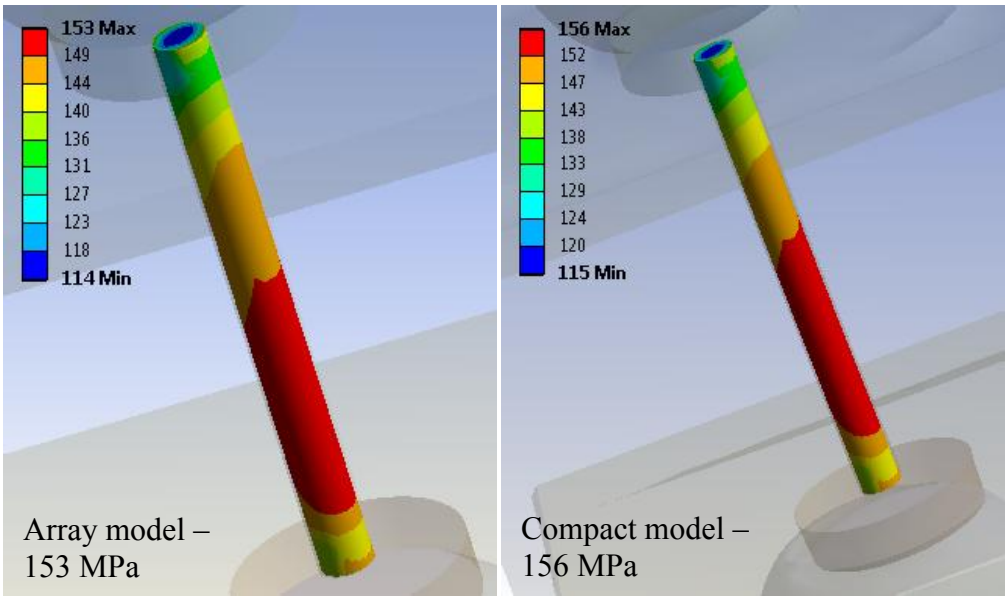


Figure 4-20 Von-mises stress in TSV (Cu region – sub model)

It is clearly demonstrated by Figure 4-15 through Figure 4-20 that the compact model prediction is in excellent agreement with the full array configuration. The magnitude and the distribution of the mechanical response (warpage, TSV region stress-strain distribution) in the global and the local model is very similar between the array and the compact model configuration with a maximum error of 7% in the principal stress in the SiO_2 dielectric layer.

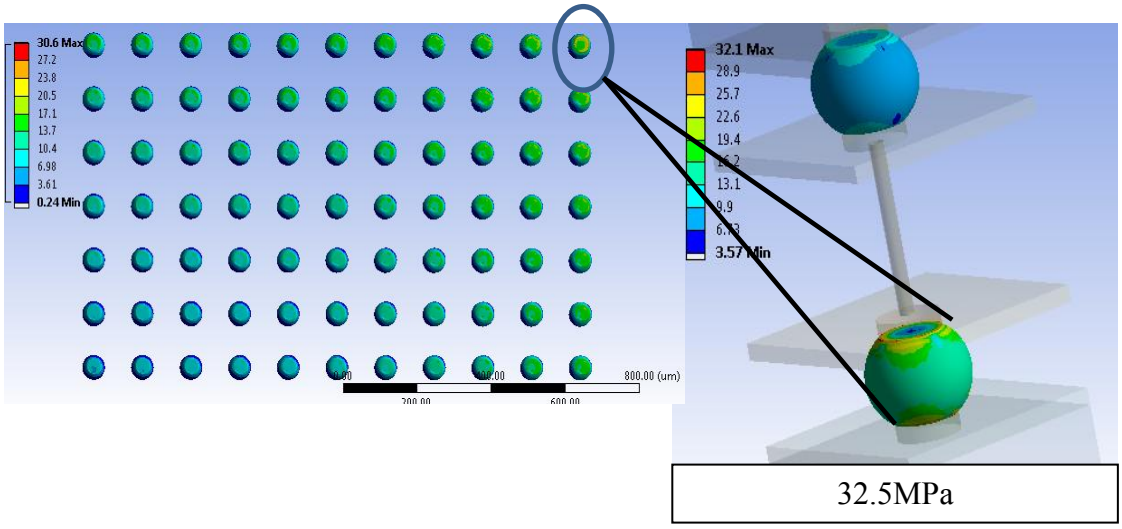


Figure 4-21 Solder μ-bumps von mises stress (Array model – MPa)

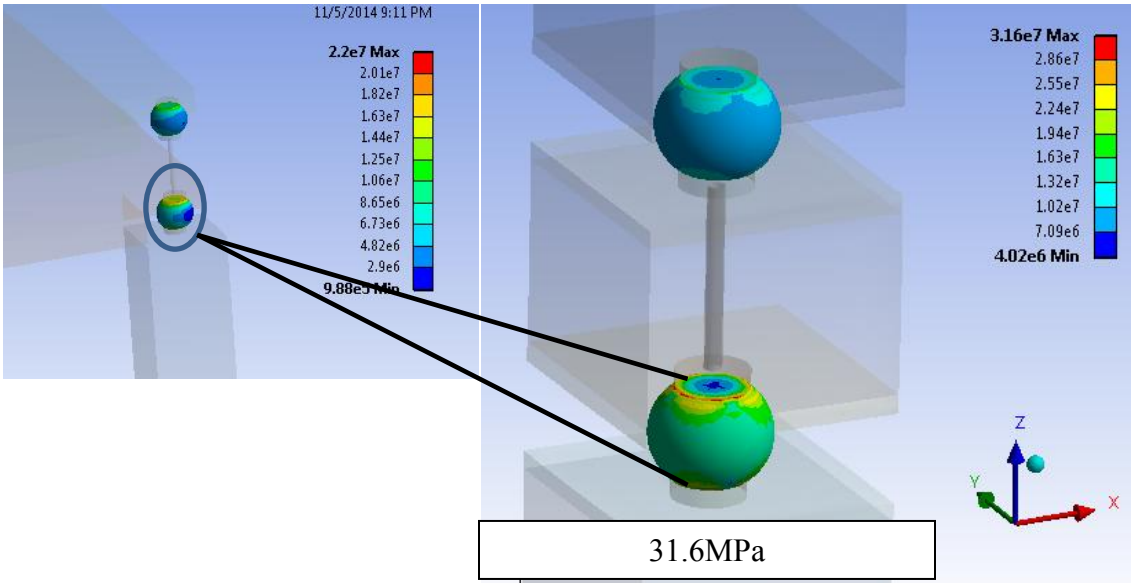


Figure 4-22 Solder μ-bumps von mises stress (compact model – MPa)

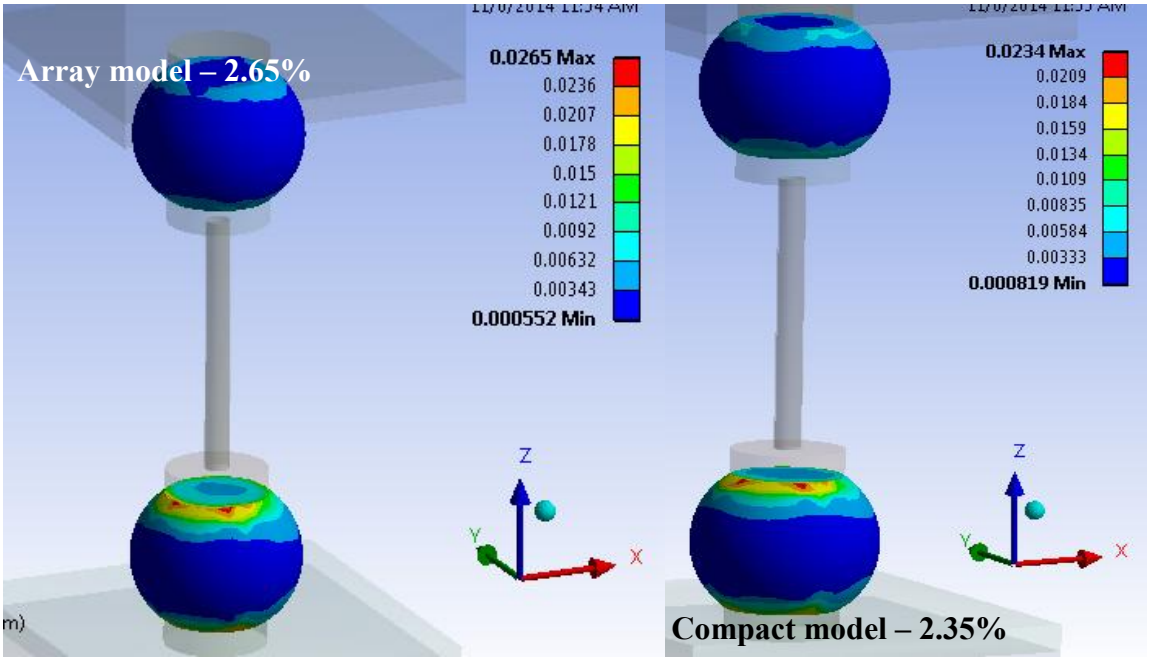


Figure 4-23 Solder µ-bumps total strain

It is clearly demonstrated by Figure 4-21 through Figure 4-23 that the compact model prediction is in good agreement with the full array configuration. Interestingly, there is a significant variation between the global von mises stress for the 2 configurations but the more accurate sub-model results are in excellent agreement. The magnitude and the distribution of the mechanical response is very similar between the 2 configurations with a maximum error of 11% in the total equivalent strain.

Table 4-3 Summary of Results

Warpage (global model) (μm)			TSV Stress – SiO2 (MPa)			Solder μ-bump stress (MPa)		
Array Model	Compact Model	% Error	Array Model	Compact Model	% Error	Array Model	Compact Model	% Error
33.2	31.6	5	224	211	6	32.1	31.6	1.6

Table 4-4 Summary of Results – strain validation

Solder μ-bump strain Local model (% strain)			TSV strain – Cu Local model (% strain)			TSV strain SiO2 Local model (% strain)		
Array Model	Compact Model	% Error	Array Model	Compact Model	% Error	Array Model	Compact Model	% Error
2.65	2.35	11	.375	0.4	2.5	0.69	0.68	1.5

Table 4-3 and Table 4-4 show the summary of the various output correlation parameters between the array and the compact model configurations. The maximum error % observed is in the solder bump von mises stress (11%) while all the other parameters depict excellent agreement.

It is to be noted that the array configuration (quarter symmetry model) with only 121 TSVs and 242 μ-bumps took about 30 hours to solve on a 32GB RAM system with high-performance-computing (HPC) capability while the compact model for a similar

footprint solves in few minutes (~100x reduction in the solution time). The reduction in the solution time is a result of replacing the TSV silicon (Si) system with an effective Si block while replacing the array of μ-bumps with a viscoplastic solder block. Now, we can see that if a similar analysis is to be done for a larger die 3-D TSV footprint with fine pitch TSVs/μ-bumps, it will be impractical or may not be even possible in some cases, thereby warranting a compact modeling approach.

4.4.2 Effect of TSV Joule Heating on Device Performance

Thermal distribution in the package is analyzed for different TSV currents including a base-line case of no-current through the TSVs and the junction temperature is determined for each case. The quarter symmetry 3-D model with a 46x46 TSV array is solved for the temperature distribution. Effective thermal block for the interconnect μ-bump is used with an effective thermal conductivity value. The response from the thermal analysis is correlated to the device performance using the correlations available in the literature.

Joule Heating: Q (Joule) = $I^2 R$; Equation 4-1

I = current through TSV (A); R=TSV Resistance (ohm)

The thermal analysis was done with the commercially available code ANSYS WORKBENCH v14.5. The test vehicle (TV) consisting of substrate, 2 stacked dies, TSVs, underfill/pillar effective block (Cu & Pb-free solder cap), and the mold was

formulated. Each die had an area of 49 mm^2 with 100 μm thickness. TSV diameter was 10 μm with 1 μm thick oxide insulation and a pitch of 75μm [35]. Substrate and the inter-die bond layers (interconnects) were modeled as effective blocks with calculated values of effective thermal conductivity [50]. This was done to achieve the desired accuracy within a reasonable computational time. A quarter symmetry global model was formulated. Adiabatic boundary condition was applied at the symmetry faces, and each die was thermally loaded with uniform power, so as to have power of 1 W in the top die ($2W/cm^2$) and 3W, 2W, 1W in the bottom die ($6W/ cm^2$, $4W/ cm^2$, $2W/ cm^2$, respectively). TSV joule heating has been modeled as internal heat generation. Thermal analysis was performed as a natural convection problem with a convective heat transfer coefficient value of $8W/m^2K$ at the exposed surfaces of the overmold. PCB has not been modeled in the analysis. To compensate for the PCB area, a higher "h" value has been used for the substrate bottom/side faces. Ambient is at 22°C. Material properties for the package components were derived from the literature and are given in Table 3 [51, 52]

A 3-D quarter symmetry model of the 3D IC with approximately 8000 TSVs is formulated in ANSYS. Different configurations (with varying TSV current) and the baseline case with no TSV current are simulated and thermal analysis is performed to determine the thermal profile in the package. The uniform chip power is 1W for the top die and 3W, 2W, and 1W for the bottom die (die with TSVs). Figures (see Figure 4-24 and Figure 4-25) demonstrate the temperature distribution in the package and the bottom die for the baseline case (no TSV current) and 40mA TSV current case, respectively.

Total package power was 4W (top die – 1W, and bottom die – 3W). Junction temperature in the bottom die is found to be 88.4°C for the baseline case while 99.3°C for the 40mA TSV current case.

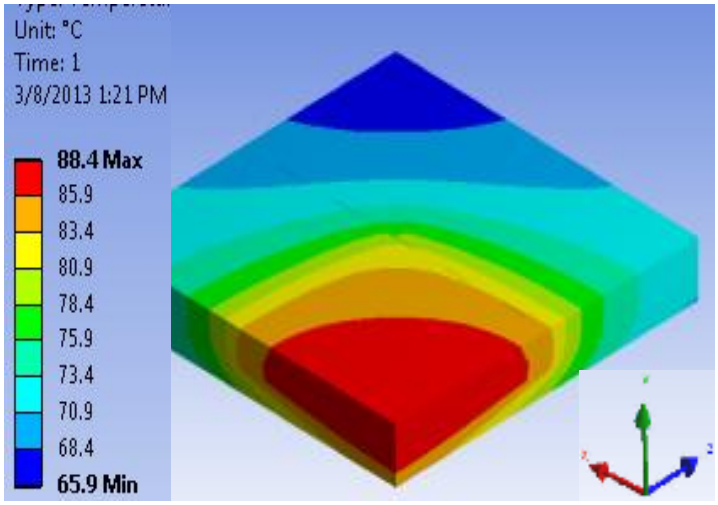


Figure 4-24 Temperature distribution (no TSV current)

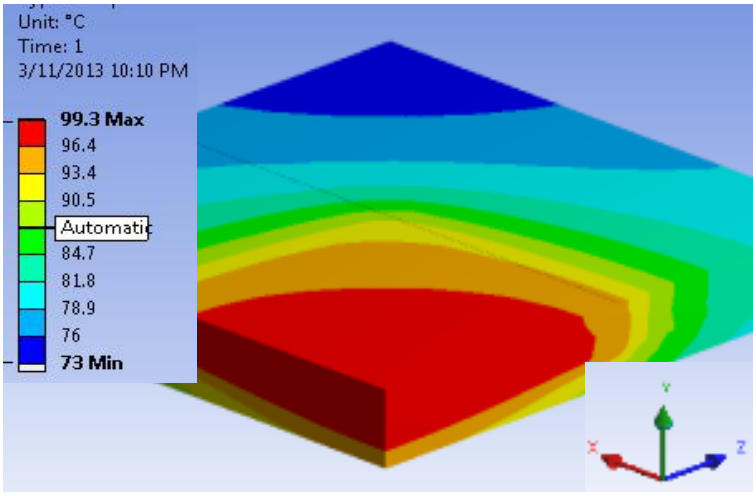


Figure 4-25 Temperature distribution

(40 mA TSV current)

Full field thermal analysis results show 11°C increment (12% increase) in the junction temperature between the baseline and the 40mA case when the total package power is 4W. Figure (see Figure 4-26) demonstrates the junction temperature variation with the TSV current for all the cases analyzed i.e., (Total power = 2W, 3W, and 4W). As expected the temperature increases with higher TSV currents and the difference was significant. From figure (see Figure 4-28), it is evident that the variation in the junction temperature with TSV current is independent of the total power of the package.

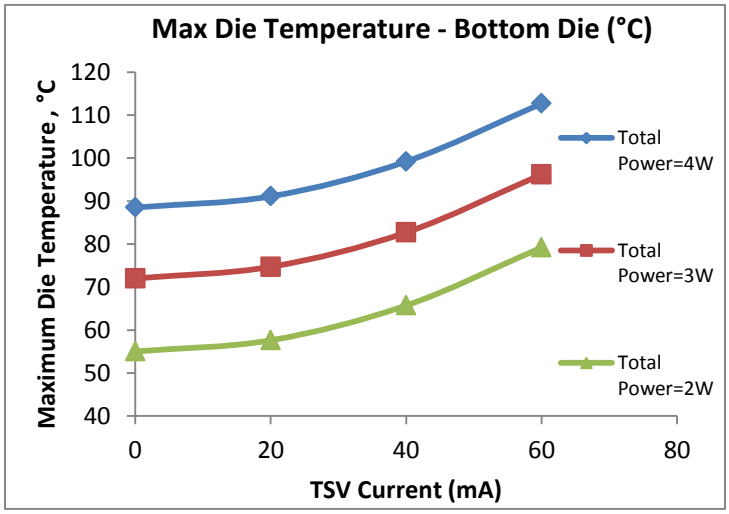


Figure 4-26 Junction temperature variation with TSV current

Using equations 4-2 through 4-4**,** the junction temperature increase with joule heating is correlated with device performance. Figure (see Figure 4-27) demonstrates the variation of the performance parameters: e- mobility, transconductance, and channel noise with the TSV current (0 mA to 60mA). Almost 10% performance hit is observed for e- mobility and transconductance and ~15% for the channel thermal noise between the baseline and

the 60mA case. Furthermore, the effect of the TSV thermal conductivity on the thermal and device performance of the 3-D TSV IC was also studied. Since TSVs occupy 1.5% of the chip real estate (CRE), the majority of the heat is conducted through the silicon (~98.5% of the chip area). Therefore, variation of the TSV thermal conductivity does not have any significant effect on the overall temperature profile of the package. The junction temperature shows negligible change when the TSV thermal conductivity is varied, thereby having no effect on the device performance. Figures (see Figure 4-29 through Figure 4-31) show that the e- mobility, transconductance, and the channel noise is independent of the TSV thermal conductivity. This tells us that TSVs primarily serve as the signal carriers and the majority of the heat dissipated is through the silicon owing to the TSVs small area (~2% of the total die area).

$$\mu(T) = \mu(T_0) \cdot (T/T_0)^{-1.5} \qquad \text{Equation 4-2}$$

$$\beta(T) = \beta(T_0) \cdot (T/T_0)^{-1.5} \qquad \text{Equation 4-3}$$

$$\sqrt{v^2} = \sqrt{8.k.T}/\{\sqrt{(3.\sqrt{(2.\beta.I)}\}} \qquad \text{Equation 4-4}$$

$\mu(T)$ = e- mobility parameter

$\beta(T)$ = Transconductance Parameter

v^2 = Channel Noise

T = Transistor Operation Temperature (°K)

I = Drain Current (A)

k = Boltzmann Constant

T_0 = Transistor Operation Temperature – base case (°K)

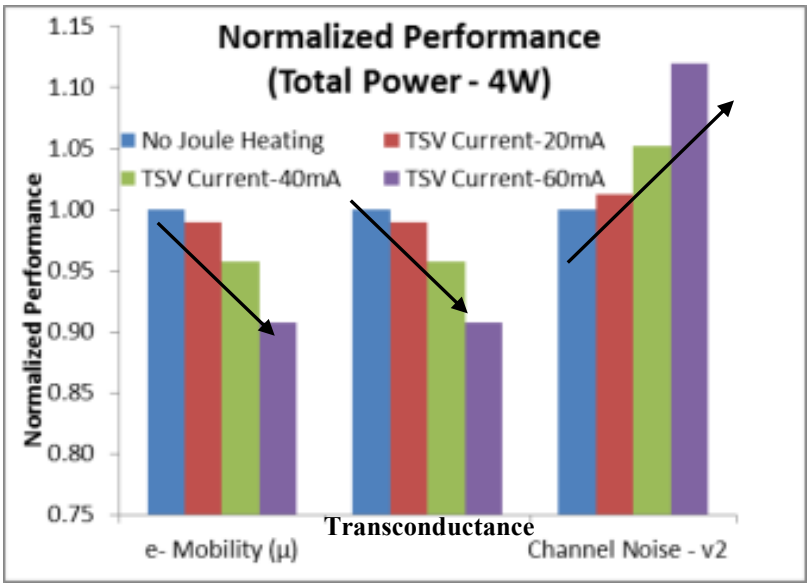


Figure 4-27 Performance degradation with TSV joule heating

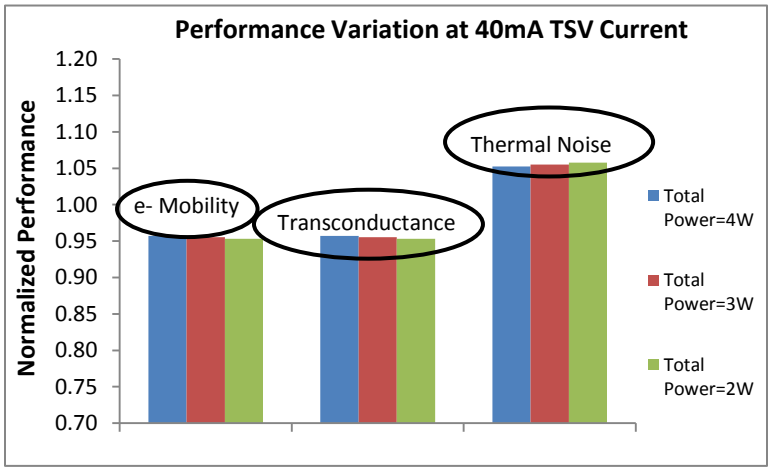


Figure 4-28 Variation in performance hit with chip power for 40mA case (normalized to the baseline case)

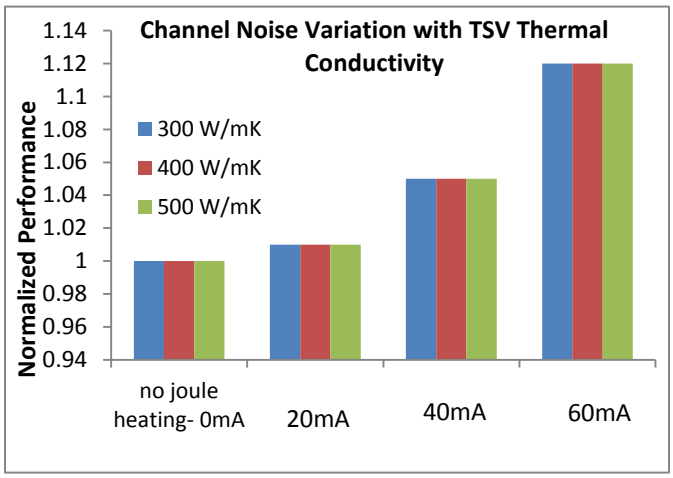


Figure 4-29 Effect of TSV thermal conductivity on channel noise

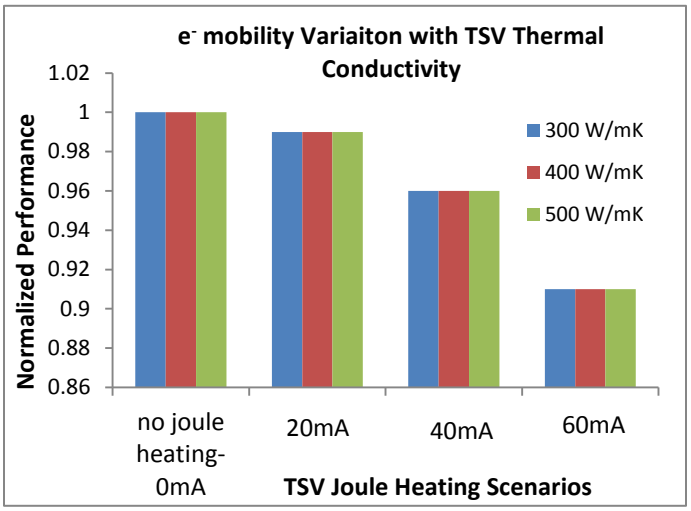


Figure 4-30 Effect of TSV thermal conductivity on e^- mobility

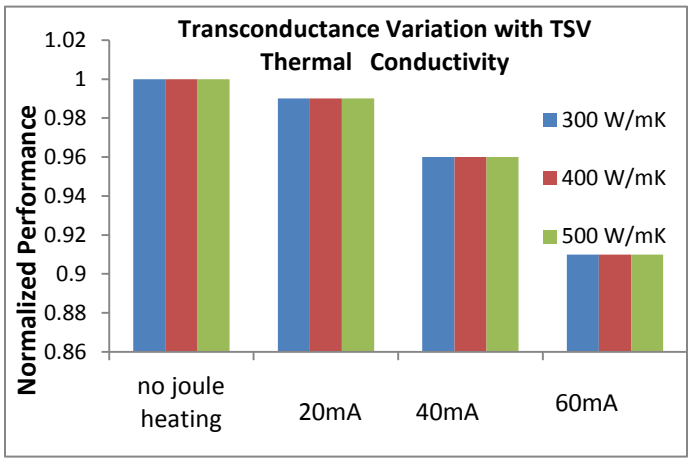


Figure 4-31 Effect of TSV thermal conductivity on transconductance

Chapter 5

SUMMARY AND CONCLUSION

A novel 3D compact modeling framework has been developed to perform multi-level chip-package-interaction (CPI) and board-level thermo-mechanical reliability analysis for different advanced packages during assembly and thermal cycling. CTE mismatch forces are highest at regions away from the die center, so a detailed model of the die corner region has been built, while a visco-plastic material model has been developed for the rest of the interconnect region (effective block) to preserve the temperature and rate dependence of the structure.

Compact model has been rigorously validated against a full-scale 3D array model for standard stress parameters. Qualitative and quantitative correlation work has been presented to give confidence over the accuracy and scalability of the proposed framework. The reduction in numerical compute time is up to 15x or more for a typical flip chip BGA configuration with little compromise in accuracy (maximum +/-15%).

The developed compact modeling methodology is scalable across different solder alloys, different types of loadings/rates and different advanced packages (SOCs and 3D ICs), thereby making it a viable technique that could be easily leveraged during development and sustainability. The developed methodology can scale across various current (28nm) and future (20nm, 20nm+) technologies to mitigate the CSAM white-bump risk during chip-package assembly. The compact model is also qualitatively validated against existing literature thus gaining confidence over its accuracy and applicability in design-for-reliability studies (DFR).

A correlation is developed between the modified ANAND's viscoplastic material model constants (s_o and S^) for the effective block and the solder volume fraction for SA305 alloy. This relation will facilitate instantaneous determination of the ANAND's viscoplastic constants for the effective block, thereby resulting in significant time reduction for tuning.

There is a significant move to introducing 3-D stacking in a variety of electronic systems. Most of the products in the market are low power devices and utilize wire-bond for interconnects. As we start migrating this technology to more general systems, the use of TSVs is critical in both increasing the number of interconnects per area (peripheral vs area) as well as improving the electrical performance. Thermal-Electrical analysis of a 3D IC with TSVs is demonstrated. It is clearly evident that joule heating has a significant contribution to the overall heat in the package, thereby causing a significant die temperature increase. Joule heating can cause a temperature increase of up to 25°C when compared with baseline (no TSV current case). A 10-15% performance hit is seen between the baseline and the 60mA case for various performance parameters. The performance hit variation (at constant TSV current) is negligible with the chip power i.e., performance hit due to joule heating is independent of chip power, and that it only depends on the TSV resistance and current. The temperature and the device performance degradation are practically independent of the TSV thermal conductivity owing to the fact that TSVs occupy a very small area on the die, and that the majority heat dissipation is through the silicon.

Bibliography

[1] Z. Or-Bach, "Is the cost reduction associated with IC scaling over?," *EE Times,* 16 July 2012.

[2] D. Agonafer, A. kaisare, M. Hossain, L. J. Yong, T. Dishongh and S. Pekin, "Thermo-Mechanical Challenges in Stacked Packaging," *Heat Transfer Engineering,* vol. 2, no. 29, pp. 134-148, 2006.

[3] K. Banerjee, S. J. Souri and K. C. Saraswat, "'3-D ICs: A novel chip design for improving deep submicron interconnect performance and systems-on-chip integration," *Proc IEEE,* vol. 5, no. 89, pp. 602-633, 2001.

[4] S. Pozder, R. Chatterji, A. Jain, Z. Huang, R. E. Jones and E. Acosta, "Progress of 3D integration technologies and 3D interconnects," in *Proc. IEEE Intl. Interconnect Tech. Conf.*, San Fransisco, 2007.

[5] S. M. Alam, R. E. Jones, S. Pozder, R. Chatterjee and A. Jain, "New design considerations for cost effective three-dimensional (3D) system integration," *IEEE Trans VLSI Systems,* vol. 18, no. 3, pp. 450-460, 2010.

[6] E. J. Vardaman, "Applications for TSV and Issues for Adoption," 2008.

[7] R. N. Master, "Packaging, Reliability and Failure Analysis in Solving Future

Packagign Challenges," in *Interpack*, 2011.

[8] G. E. Moore, "Cramming more Components onto Integrated Circuits," *Electronics,* vol. 8, no. 38, 1965.

[9] A. Wolfgang, M. Brillouët, P. Cogez, M. Graef, B. Huizing and . R. Mahnkopf, "More-Than-Moore," *White paper.*

[10] "The International Technology Roadmap for Semiconductors,," *Semiconductor Industry Association,* no. Edition. 5 , 2009.

[11] "wikipedia," wikipedia, [Online]. Available: http://en.wikipedia.org/wiki/System_on_a_chip.

[12] "The Great Debate: SOC vs. SIP," August 2009. [Online]. Available: http://www.eetimes.com/document.asp?doc_id=1153043. [Accessed 15 August 2014].

[13] "Apple Launches New iPad," Apple, 7 March 2012. [Online]. Available: www.wikipedia.org.

[14] "Apple Press Info," 23 October 2012. [Online]. Available: http://www.apple.com/pr/library/2012/10/23Apple-Introduces-iPad-mini.html. [Accessed 15 August 2014].

[15] D. Agonafer and B. Sammakia, "Challenges in 3D Packaging," in *InterPACK*, Vancouver BC, 2007.

[16] Z. Or-Bach, "Monolithic 3D is Now in Production: Samsung Starts Mass Producing Industry's First 3D Vertical NAND Flash," MonolithIC3D, 19 August 2013. [Online]. Available: http://www.monolithic3d.com/blog/monolithic-3d-is-now-in-production-samsung-starts-mass-producing-industrys-first-3d-vertical-nand-flash. [Accessed 16 August 2014].

[17] "A revolution in Memory," MIcron, 2 April 2013. [Online]. Available: http://www.micron.com/products/hybrid-memory-cube/all-about-hmc. [Accessed 16 August 2014].

[18] Venkatadri, B. Sammakia, K. Srihari and D. Santos, "A Review of Recent Advances in Thermal Management in Three-Dimensional Chip Stacks in Electronic Systems", Journal OF Electronics Packaging," *Journal of Electronics Packaging,* vol. 133, pp. 1-15.

[19] J. Zheng, Z. Zhang, Y. Chen and J. Shi, "3D Stacked Package Technology and its Application Prospects," in *International Conference on New Trends in Information and Service Sciences* , 2009.

[20] Ouyang, E; MyoungSu, C; Chow, S; Emigh, R; Joshi, M; Martin, R; Pendse, R, "Improvement of ELK Reliability in Flip Chip Packages using Bond-on-Lead (BOL)

Interconnect Structure," in *IMAPS*, 2010.

[21] F. Kuechenmeiste, "Chip Package Interaction: Challenges and Solutions to Mechanical Stability of BEOL at 28nm Node and Beyond for Advanced Flip Chip Application," in *Electronics Packaging Technology Conference*, December 2012.

[22] C. Shah, A. Karmarkar and X. Xu, "Modeling of Interconnect Stress Evolution during BEOL Process and Packaging," in *International Interconnect Technology Conference, IITC*, June 2013..

[23] C. Liu and K. Chiang, "“Solder shape design and thermal stress/strain analysis of flip chip pacake usign hybrid method," in *Proc. 2000 Int. Symp. Electronic Materials Packaging*, Hongkong, 2000.

[24] K. Chiang, Y. Lin and H. Cheng, "On enhancing eutectic solder joint refliability using a second-reflow process approach," *IEEE Transactions of Components and Paackaging, Manufacturing Technologies,* vol. 23, pp. 9-14, 2000.

[25] S. F. Popelar, "A parametric study of flip chip reliability based on solder fatigue," in *Proc. IEEE Components Manufacturing Packaging*, 1997.

[26] J. T. A. Z. X. K. V. L. Y. Y. D. C. C. T. J. H. C. S. D. Ong, "Optimization of the Thermo-mechanical Reliability of a 65 Cu/low-k Large-Die Flip Chip Package," *IEEE Transactions on Components and Packaging Technologies,* vol. 32, no. 4, pp.

838-848, 2009.

[27] J. C. D. Pang, "Flip Chip on Board Solder Joint Reliability Analysis Using 2-D and 3-D FEA Models," *IEEE Transactions on Advanced Packaging,* vol. Vol. 24, no. 4, pp. 499-506, 2001.

[28] X. Chen, G. Chen and M. Sakane, "Prediction of Stress-Strain Relationship with an Improved Anand Constitutive Model for Lead Free Solder Sn-3.5Ag," *IEEE Transactions on Components and Packaging Technologie,* vol. Vol. 28, no. 1, pp. 11-116, 2005.

[29] Y.-L. T. T. C. R. S. G. L. M. K. H. P. H. P. R. H. K. W. Larry Lin, "Second-Level Interconnects Reliability for Large-die Flip Chip Lead-Free BGA Package," in *IEEE Electronic Components and Technology Conference*, 2011.

[30] L. L. Mercado, V. Sarihan, Y. Guo and A. Mawer, "Impact of Solder Pad Size on Solder Joint Reliability," *IEEE TRANSACTIONS ON ADVANCED PACKAGING,* vol. 23, no. 3, pp. 415-420, 2000.

[31] Y. Lai and T. Wang, "Verification of Sub Modeling Technique in Thermo Mechanical Reliability Assessment of Flip-chip Package Assembly," *Microelectronics Reliability,* 2005.

[32] H. C. Cheng, K. N. Chiang and C. K. Chen, "Parametric Analysis Of Thermally

Enhanced Bga Reliability Using A Finite-Volume-Weighted Averaging Technique," in *IMECE*, 1998.

[33] G. G. I. K. V. M. E. Gustafsson, "Finite Element Modeling of BGA Packages for Life Predictions," in *Eletronic Components and Technology Conference*, 2000.

[34] C. Shah, F. Mirza and C. S. Premachandran, "Chip Package Interaction(CPI) Risk Assessment On 28nm Back End Of Line(BEOL) Stack Of A Large I/O Chip Using Compact 3D FEA Modeling," in *EPTC*, Singapore, 2013.

[35] Ku, T, "3D TSV Stacking IC Technologies, Challenges and Oppurtunities," ITRI , 2011.

[36] A. Jain, R. E. Jones, R. Chatterji, S. Pozder, Huang and Z. Huang, "Analytical and numerical modeling of the thermal performance of three-dimensional integrated circuits," *IEEE Trans. Components Packaging Technol.,* vol. 33, no. 1, pp. 56-63, 2010.

[37] C. Selvanayagam, J. Lau, X. Zhang, S. Seah, K. Vaidyanathan and T. Chai, "Nonlinear Thermal Stress/Strain Analyses of Copper Filled TSV (Through Silicon Via) andTheir Flip-Chip Microbumps," *IEEE TRANSACTIONS ON ADVANCED PACKAGING,* vol. 32, no. 4 , pp. 720-728, NOVEMBER 2009.

[38] *A Model for the Free (Top) Surface Deformation of Through-Silicon Vias.*

[39] J. Kawa, "TSV Stress Management," Synopsys, 2010.

[40] S. Moon, S. Prstic and C. P. Chiu, "Thermal Management Of A Stacked-Die Package In A Handheld Electronic Device Using Passive Solutions," *IEEE Journal,* pp. 791-797, 2006.

[41] S. P. Tan, X. W. Zhang and D. Pinjala, "Prediction of Hotspots on 3D Packages due to Joule Heating in Through Silicon Vias (TSV)," in *EPTC*, 2009.

[42] S. Bansal, B. Griffin and M. Greenberg, "3-D IC design: New possibilities for the wireless market," *EE Times,* 7 June 2011.

[43] B. Sung, K. V. Chepuri, D. Agonafer and A. Cristina, "Thermal Enhancement of Stacked Dies Using Thermal Vias," in *" International Mechanical Engineering Congress and Exposition*, Chicago, 2006.

[44] K. Matsumoto, S. Ibaraki, K. Sakuma, Y. Oril and F. Yamada, "Investigation of Cooling solutions for Three-dimensional (3D) Chip Stacks," in *26th IEEE SEMI-THERM Symposium*, 2010.

[45] K. Kota, P. Hidalgo, Y. Joshi and A. Glezer, "A Novel Conduction-Convection based cooling Solution for 3D Stacked Electronics," in *26th IEEE SEMI-THERM Symposium*, 2010.

[46] H. Phan and D. Agonafer, "Experimental Analysis Model of an Active Cooling method for 3D-ICs Utilizing a Multidimensional Configured Thermoelectric," in *26th IEEE SEMI-THERM Symposium*, 2010.

[47] D. Agonafer and H. Phan, "Multidimensional heat transfer system for cooling electronic component". USA Patent US20120211204-A1., 23 August 2012.

[48] S. Karajgikar, D. Agonafer, K. Ghose, B. Sammakia and C. Amon, "Multi-Objective Optimization to Improve Both Thermal and Device Performance of a Non-uniformly Powered Micro-Architecture," *Journal Of Electronics Packaging,,* vol. Vol 132., June 2010.

[49] "colorado.edu," [Online]. Available: http://ecee.colorado.edu/~bart/book/transpor.htm, accessed 8/12/2014.

[50] F. Mirza, T. Raman, A. Iftakher, S. Iqbal and D. Agonafer, "Parametric Thermal Analysis of TSVs in a 3-D Module based on Interconnect Delay and Silicon Efficiency," in *ITherm*, 2013.

[51] A. Menon, S. Karajgikar and D. Agonafer, "Thermally design optimization of a package on package," in *5th IEEE SEMI-THERM Symposium*, 2009.

[52] Y. T. A. A. N. T. F. F. T. Abe, "High Thermally Conductive and High Reliability Underfill," in *Proceedings of HDP* , 2006..

[53] "EE 8843 Analog IC Design," [Online]. Available: http://www.coursehero.com/file/1710655/pp39-46. [Accessed 12 August 2014].

[55] [Online]. Available: http://www.future-fab.com/assets/images/ff17_wp_chandran_fig2.gif. [Accessed 16 August 2014].

[56] [Online]. Available: http://www.future-fab.com/assets/images/ff17_wp_chandran_fig2.gif.

[57] C. Lee, "The trend of TSV packaging", 2011

[58] S.B. Park, I. Ahmed, 'Shorter Field Life in Power Cycling for Organic Packages" ASME Journal of Electronics Packaging, 2007, Volume 129, pp.28-34

[59] X. Liu, "Failure Mechanisms and Optimum Design for Electroplated Copper Through-Silicon Vias (TSV), ECTC 2009, pp.624-629

[60] Fan, X.J., Han, Q., Varia, B., "Effects of Design, Structure and Material on Thermal-Mechanical Reliability of Large Array Wafer Level Packages, 2009 International Conference on Electronic Packaging Technology & High Density Packaging (ICEPT-HDP)

[61] Tee, T., Ng, H., Yap, D., Baraton, X., Zhong, Z., "Board level solder joint reliability modeling and testing of TFBGA packages for telecommunication applications", Micro-electronics Reliability, 2003, pp. 1117-1123

[62] R. Darveaux, "Effect of simulation methodology on solder joint crack growth correlation," in *ECTC*, Las Vegas, 2000.

[63] L.F. Coffin, "A study of the effects of cyclic thermal stresses on a ductile metal," ASME, 1954"

[64] S.S. Manson, "Fatigue: a complex subject-some simple approximation," Exp Mech 5(7):193–226, 1965.

[65] A. Syed, "Accumulated creep strain and energy density based thermal fatigue life prediction models for SnAgCu solder joints," in ECTC, 2004, pp. 737-746

Biographical Statement

Fahad Mirza received the Master's degree (MS) in mechanical engineering from SUNY Binghamton, NY where his research focused on high-strain rate reliability analysis and development of an impact reliability correlation model for Pb-free micro-electronics.

Fahad is a Graduate Research Assistant (GRA) in the Electronic, MEMS and Nano-electronics Packaging Center (EMNSPC) at the University of Texas at Arlington. His current research focuses on the development of a compact modeling technique for multi-level thermo-mechanical reliability analysis. During his PhD, Fahad iinterned at GLOBALFOUNDRIES, NY in 2012-2013 where he worked on advanced technology node development and 3D packaging. He has over 20 technical publications including journal articles and conference papers of international repute.