Effect of Finite Element Modeling Techniques on Solder Joint Fatigue Life Prediction of Flip-Chip BGA Packages

Xuejun Fan1, Min Pei2, and Pardeep K. Bhatti1
1Intel Corporation, M/S CH5-263, 5000 W Chandler Blvd, Chandler, AZ 85226
2Georgia Tech, 801 Ferst Dr. NW, Atlanta, GA 30332-0405
xuejun.fan@intel.com

Abstract

Solder joint fatigue life in thermal cycling has been studied for decades using the finite element method. A great variety of modeling methodologies such as global/local modeling (sub-modeling) and sub-structure modeling (superelement) has been developed. Many different types of constitutive equations for solder alloys, various loading assumptions, and several definitions of damage parameters have been used. However, the accuracy of these different modeling approaches has not been completely evaluated in literature. There has been some long-standing confusion regarding the modeling assumptions and their effect on the accuracy of models, such as the initial stress-free temperature setting, selection of damage parameters, and choice of element type. This paper presents a comprehensive study of finite element modeling techniques for solder joint fatigue life prediction. Several guidelines are recommended to obtain consistent and accurate finite element results.

Introduction

Finite element method has been used for a long time to study the solder joint fatigue life in thermal cycling. The fatigue modeling process consists of four primary steps [1]. First, the constitutive material models are chosen. Selection of an appropriate constitutive model that describes the solder behavior in the actual packaging application is critical to obtain accurate results. Second, the FEA model is created with the appropriate boundary conditions. The stress/strain values are calculated in this step. Third, the FEA results are used to develop a fatigue equation that predicts the number of cycles to failure, Nf. The form of fatigue model depends on the constitutive model selected in the first step. For example, if damage mechanics approach is used, the fatigue model might be based on void coalescence and growth. Fourth, the model results must be verified using thermal cycling test data. These four steps describe the general process by which a predictive fatigue model is developed and verified.

Available solder joint fatigue models can be categorized as (i) strain based [2][3][6], (ii) strain energy based [4][5][6], (iii) fracture based [7][8][9], and (iv) damage based [10]. Stress based models are generally not used because low cycle fatigue is the dominant solder joint failure mode in thermal cycling. But the stress based model can be potentially used at shock and vibration load conditions.

This paper will focus on the second step - the creation of finite element model and the extraction of damage parameters, by investigating the effect of modeling techniques on the accuracy of board-level solder joint fatigue life prediction. A flip-chip ball grid array (BGA), commonly used packaging technology in computer and many other applications, is considered in this investigation. Solder creep is assumed to be the dominant material behavior for both SnPb and SnAgCu materials. A double-power creep law is used [11][12] in this paper. A full non-linear finite element model, in which several solder balls at the locations of interest have a refined mesh pattern, is created as a baseline to compare and investigate the accuracy of global/local and sub-structure modeling methods. The worst case solder joint location is then identified based on the full non-linear model results and by using various damage metrics. The effect of initial stress-free condition in the simulation on model results is investigated. The choice of element type for PCB and substrate to accurately describe the bending behavior of package is discussed in detail. The effect of tie constraints (or multi-point constraints) is also investigated. The results show that very consistent finite element results can be obtained when certain rules are followed.

Finite Element Modeling

The flip-chip BGA package is modeled using the commercially available finite element software ABAQUS. A Python language script was written to generate BGA finite element models automatically with various geometric parameters.

In this paper, only the second level interconnect (solder joint between the package and the PCB) is studied. A full non-linear, quarter symmetry finite element model was created. In this model, half of the solder joints under the die shadow region have a refined mesh pattern, as shown in Figure 1.. The rest of the solder joints are modeled with a relatively coarse mesh. Cross sections of a coarse and a refined cell are shown in Figure 1(b) and Figure (c), respectively. Copper pads on each side of the solder joint are modeled in both cases. The solder joint is solder mask defined (SMD) on the package side and metal defined (MD) on the board side. Details of both coarse and refined mesh patterns and geometry of solder balls are shown in Fig. 2. The refined solder ball cell uses mesh transition to connect to the coarse mesh without the use of tie or multi-point constraints.

Material Properties

There is considerable variation in the published test data on mechanical properties of solder alloys due to tolerances in the measurement equipment/techniques and variability in test specimen design and preparation. This has resulted in several constitutive material models to describe the plasticity and steady state creep behavior of solders. One challenge has been how to separate the time-independent (plastic) and time-dependent (creep) inelastic components from the measured strains, especially at high temperatures. A combined creep and plasticity material model, which captures the total strain behavior in the operating range, was proposed by Wong, Helling, and Clark for 63Sn37Pb eutectic alloy [12]. Bhatti et
al. [13][14][15] implemented this constitutive model and developed 3-dimensional package level finite element models to perform solder joint creep simulations. This material model can be written as

$$\dot{\varepsilon} = \frac{\dot{\sigma}}{E} + B_1 D \left( \frac{\sigma}{E} \right)^3 + B_2 D \left( \frac{\sigma}{E} \right)^7$$

(1)

where

- $\dot{\varepsilon}$ = Total strain rate (1/sec)
- $\sigma$ = Stress (MPa)
- $E$ = Modulus of Elasticity (MPa) = 56000 – 88T
- $T$ = Temperature (K)
- $B_1 = 1.70 \times 10^{12}$ 1/sec
- $B_2 = 8.90 \times 10^{24}$ 1/sec
- $D = \exp \left( - \frac{5413}{T} \right)$

The second term in the equation (1) accounts for the grain boundary sliding (GBS) creep strain and the third term accounts for the matrix creep (MC) strain.

Wiese et al. [11] studied the creep behavior of bulk, PCB sample, and flip chip solder joint samples of Sn/4.0Ag/0.5Cu solder and identified two mechanisms for steady state creep deformation for the bulk and PCB samples. They attributed these to climb controlled (low stress) and combined glide/climb (high stress) mechanisms and represented steady state creep behavior using a double power law model as shown below

$$\dot{\varepsilon} = \frac{\dot{\sigma}}{E} + A_1 D_1 \left( \frac{\sigma}{\sigma_n} \right)^3 + A_2 D_2 \left( \frac{\sigma}{\sigma_n} \right)^{12}$$

(2)

where

- $\dot{\varepsilon}$ = Total strain rate (1/sec)
- $\sigma$ = Stress (MPa)
- $E$ = Elastic Modulus (MPa) = 59533 – 66.667T
- $T$ = Temperature (K)
- $A_1 = 4.0 \times 10^{-7}$ 1/sec
- $A_2 = 1.0 \times 10^{-12}$ 1/sec
- $D_1 = \exp \left( - \frac{3223}{T} \right)$
- $D_2 = \exp \left( - \frac{7348}{T} \right)$
- $\sigma_n = 1$ MPa

The second term in the equation (2) represents the climb controlled creep strain and the third term represents the combined glide/climb strain. Syed [6] applied this creep model to develop a fatigue life model for SnAgCu solders.

Published material properties [6] are used for all other materials as listed in Table 1.

### Table 1 Material Properties

<table>
<thead>
<tr>
<th>Material</th>
<th>Young's Modulus (GPa)</th>
<th>Poisson's Ratio</th>
<th>Coefficient of Thermal Expansion (ppm/°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silicon</td>
<td>131.0</td>
<td>0.3</td>
<td>2.6</td>
</tr>
<tr>
<td>Copper</td>
<td>128.7</td>
<td>0.3435</td>
<td>17.0</td>
</tr>
<tr>
<td>Underfill</td>
<td>9.9</td>
<td>0.23</td>
<td>24.0</td>
</tr>
<tr>
<td>Substrate</td>
<td>22.0</td>
<td>0.11</td>
<td>17.0</td>
</tr>
<tr>
<td>PCB</td>
<td>24.2</td>
<td>0.11</td>
<td>19.6</td>
</tr>
<tr>
<td>SnPb</td>
<td>25.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SnAgCu</td>
<td>20.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Edge Singularity and Volume Averaging**

Mesh density sensitivity is a critical issue in FEA simulations with dissimilar materials. This arises from the edge singularity at the solder joint to copper pad interface. Therefore the maximum stress/strain value in the solder joint is dependent on the mesh density. In order to minimize the effect of this singularity, a widely-used technique is volume averaging over a thin layer of solder material along the solder joint and copper pad interface [6][7]. Syed used a thickness of 25 micron (1mil) for this layer in his creep strain based fatigue model [6]. Darveaux used a 30 micron thick layer in his fracture mechanics based model [7]. In this paper, 25 micron thick layers are used on the both sides of solder ball with refined mesh pattern, and modeled with two layers of elements across the thickness, as shown in Figure 2(b). Model results are averaged over these layers of elements. These include cumulated equivalent creep strain (denoted by CEEQ), cumulated creep strain energy density (ECDDEN) and Von Mises stress.

![Figure 1: A full non-linear model of a FC-BGA package, (a) quarter package, (b) coarse cell, and (c) refined cell](image_url)
Thermal Cycle Loading

Thermal cycle profile used in the paper has a range of -25 to 100°C. Dwell time at each extreme is 15min, and up / down ramp time is 8 minutes.

Accuracy of Global/Local Modeling

Although the global/local modeling approach is widely used to capture the local solder joint behavior, the accuracy of global/local modeling (also known as sub-modeling) has not been fully evaluated to the authors’ best knowledge. In this section, the sub-modeling results are compared to the results from the full non-linear model shown in Fig. 1. An example of a global/local model is shown in Figure 3. The size of the local model is equal to one solder ball pitch.

In Figure 4(a), per-cycle CEEQ (both averaged and maximum) is compared for two modeling approaches. Figure 4(b) shows the peak Von Mises stress (both averaged and maximum) during the thermal cycle, which occurs at the beginning of low temperature dwell. These results show that the difference between the two models is only 1.5% for averaged per-cycle CEEQ and averaged stress (same conclusions can be made for the averaged per-cycle strain energy density). The maximum values of per-cycle CEEQ and per-cycle creep strain energy density between the two approaches differ by up to 7%. These values are calculated at the solder ball under die shadow corner. For other solder balls under the die shadow region, this relative error can be as high as 20%, while the averaged values are always within 5% for these two approaches.

This shows that the maximum value is highly sensitive to modeling technique. However, the global/local modeling approach provides satisfactory results for the volumetrically averaged values for this package.

Accuracy of Sub-structure Modeling

Sub-structuring is a procedure that condenses a group of finite elements into one element represented as a stiffness matrix. This condensed sub-structure is called a superelement, while the rest of the structure is called the residual model. In a nonlinear analysis, one can sub-structure part of the model so that the element matrices for that portion need not be recalculated in every iteration. This approach requires only one analysis of the superelement for a unit loading condition (e.g. 1 degree rise or fall in temperature). Using appropriate scaling factors, the superelement can then be used repeatedly in the analysis of the residual structure. The limitation of using this method is that the superelement cannot include materials with non-linear, temperature dependent behavior. This limitation, however, can be overcome by either excluding these materials from the superelement or by doing multiple sub-structure analyses. The superelement approach was first used by Bhatti et al. to simulate electronic packages [14]. In order to achieve significant savings in computational time, they also included non-critical solder joints in the superelement by assuming linear behavior and temperature independent elastic modulus for these joints. The modulus of elasticity of non-critical solder joints is an artificially lowered, calibrated value of 1 GPa, which was obtained by comparing superelement and full
model results of a leaded package (e.g. PLCC and QFP) that resulted in less than 5% effect on the creep strain in the critical joint [14]. It is important to note that at the time this modeling approach was first published in 1993 by Bhatti et al. [14], available computing power made application of this technique a necessity so that 3-dimensional simulations could be completed in a practically reasonable time. However, with the computing power available in current high-end workstations, this technique is no longer required for a reasonable size package model.

In this study, the superelement approach is investigated for flip chip BGA packages and compared with full non-linear model. To save model creation time, an alternative approach equivalent to the sub-structure method is used. In this method, the full finite element model is built and only critical solder joints that would be in the residual structure are assigned non-linear creep properties. The non-critical solder balls that would be part of the superelement use linear material properties with an effective modulus of elasticity of 1 GPa. We validated that this alternate approach provides results identical to the superelement approach by running a verification superelement case.

A comparison between superelement and full nonlinear model results is shown in Figure 5(a). The results show that the superelement method gives about 60% higher averaged per-cycle CEEQ compared to the full model for SnPb. The situation is worse for the stiffer SnAgCu solder. It’s also noted that the maximum averaged CEEQ and Von Mises stress are in different solder joints in these 2 models. The reason for this difference is that the solder balls using the effective modulus of elasticity (1 GPa) do not provide the same support as real solder joints in the BGA package. The whole package becomes more compliant, and thus more deformation is introduced on solder joints. Because the effective modulus of elasticity is same for both solder materials, the stiffer SnAgCu solder introduces an even larger error. However, it is important to note that since the fatigue life equations published in [16][17] are developed based on the superelement (not the full nonlinear model) results, these models still provide very good accuracy in predicting the fatigue life for a variety of packages.

Interestingly, we also discovered that for this BGA package, when we modeled all non-critical solder joints with real temperature-dependant Young’s modulus but still without creep properties, the results were close to the full model approach, as shown in Figure 5(b). However this approach does not save significant computational time over the full nonlinear model.

In summary, sub-structure approach is very advantageous when most of the materials in a structure can be simulated with linear, temperature independent properties without a significant loss of accuracy. However, caution must be used when assumption of linearization is made.

![Figure 5: Superelement vs. Full model Results, (a) Use effective elastic modulus (1GPa) for non-critical solder joints, (b) Use temp-dependent elastic modulus (w/o creep) for non-critical solder joints](attachment:image.png)
From Figure 6(a) the solder joint in the middle of die edge (location 1-7) has the highest per-cycle average CEEQ. All solder joints along the die edge have comparable values within 20% of each other. However, the maximum per-cycle CEEQ, in Figure 6(b), shows the worst-case solder joint at the corner of die shadow (location 7-7), and the second highest value appears one row inside from the die shadow corner (location 6-6). All solder joints along die edge have relatively high strain accumulation.

The averaged and maximum Von Mises stress at the beginning of low-temperature dwell are shown in Figure 7, and show trend similar to CEEQ.

Following the analysis of Modi et al. [19], the averaged ‘peel’ stress is also plotted in Figure 8. Peel stress is defined as stress in the out-of-plane direction at the solder joint to copper pad interface. It can be seen that the solder joint one row inside from the die shadow corner (location 6-6) has the highest tensile stress. The solder joints under die corner and in the middle of die edge have compressive stress. The averaged hydraulic stress distribution, in Figure 9, shows trend similar to the peel stress.
This analysis shows that different parameters give different locations for the worst-case solder joint. Averaged per-cycle creep strain and stress shows the worst-case solder joint at the middle of die edge, while the maximum value of these two parameters show the worst joint at the die corner. Peel stress and hydraulic stress find the maximum tensile stress at the joint one row inside from the die shadow corner.

External [19] and Intel experimental data suggest that the solder joint one row inside from the corner of die shadow usually has the highest crack growth rate, and all solder joints along the die edge have comparable crack growth rates. Based on these observations, we suggest that a combination of peel stress and cumulative creep strain be used to determine the worst case solder joint location and the averaged creep strain at that location be used to develop the fatigue life model. It should be noted that this conclusion is for ‘no-preload’ case, when there is no load exerted on the package by a heatsink or another cooling device. For an investigation of the effect of preload, the reader is referred to another paper by the authors [20].

**Initial Stress-Free Condition**

There has been a long-running debate on the selection of initial stress-free temperature in solder joint modeling [21]. There are three most commonly used initial stress-free temperatures. One is the solidus temperature of solder alloy (e.g. 217°C for SnAgCu). This condition considers that the solder joints start to provide mechanical support as soon as the solder material solidifies during the reflow process. The second one is the room temperature as the initial stress-free condition (e.g. 25°C). This assumes that the shipping and storage time is sufficient to relax all the residual stresses in solder joints from the assembly process. The last one uses the high dwell temperature of thermal cycle or operating conditions (denoted as Tmax, e.g. =125°C for thermal cycling from -25°C to 125°C). This assumes that after several thermal cycles, the package reaches a stabilized cyclic pattern where the lowest stresses are seen at the end of the high temperature dwell period.

All three scenarios were simulated and the results are discussed below. For each case, a total of 10 thermal cycles were run in order to ensure that a stable cyclic pattern was achieved.

Figure 10(a) plots the history of per-cycle CEEQ for SnPb. It can be seen that the initial stress-free condition affects the per-cycle CEEQ in first two cycles of simulation only. After that the per-cycle CEEQ stabilizes quickly for all three cases and the results converge to same value. Figure 11(a) shows the peak Von Mises stress (averaged) in each cycle and Figure 11(c) plots the Von Mises stress history for all three cases. Regardless of the initial stress-free temperature, the package always relaxes to the lowest stress during the high temperature dwell period after the first cycle. This implies that whatever the initial stress state is, because of the viscous behavior of SnPb material, solder joints always experience highest stress at the beginning of the low temperature dwell period and tend to relax to ‘near-zero’ stress at the end of high temperature dwell period. The transition of ‘stress-free’ from initial condition to high-dwell temperature is completed after only one thermal cycle for SnPb. Figure 10 also shows that when the initial stress-free is assumed as Tmax or room temperature, the stabilized per-cycle CEEQ can be achieved even after the first cycle, which can save significant amount of computational time.

For SnAgCu, the history plot of the per-cycle CEEQ up to 10 cycles is shown in Figure 10(b). After 10 cycles, all three cases almost converge to same value. Unlike SnPb, more cycles are needed to achieve stabilized results if the reflow temperature is used as initial stress-free condition. This is further illustrated by the Von Mises stress plot in Figure 11(b) and (d). The peak stress shown in Figure 11(b) and the history plot in Figure 11(d) show that the stress is not fully stabilized even after 10 cycles. Again, regardless of the initial stress-free condition, the package has higher stresses at low temperature dwell and relaxes significantly at high temperature dwell period. Since SnAgCu is more creep resistant and stiffer than SnPb, it takes more cycles for SnAgCu solder joints to stabilize.
Figure 11: Averaged Von Mises stress

Figure 12 (a) and (b) show the peel stress distribution for SnPb during the 10th cycle across all solder joints under the die shadow, with the initial stress-free temperature as reflow temperature and room temperature, respectively. The trend is similar for these two cases but the values are slightly different.

This analysis shows that for viscous materials such as solder, regardless of initial stress-free condition, the structure adjusts the stress state during thermal cycling and reaches the lowest stress at the end of high temperature (Tmax) dwell. Usually the per-cycle stabilized values of strain or strain energy density are used in fatigue life prediction, therefore it is recommended to use Tmax as initial stress-free condition to achieve the stabilized solutions as quickly as possible. This increases the computational efficiency significantly.

Choice of Element Type

Hexahedral three dimensional finite elements come in linear or quadratic formulation. Furthermore, the analyst has a choice of full integration, reduced integration, incompatible mode, and hybrid element formulations. For a large BGA package, the deformation during temperature cycling is dominated by the bending of the PCB and BGA substrate. Therefore, elements that accurately model bending behavior are optimal choice for solder joint fatigue analysis. Some solid elements perform poorly in bending because of shear locking and/or hourglassing phenomena. Shear locking, or parasitic shear, is caused by an accuracy in the displacement field of a linear hexahedral element.

Figure 13 compares the CEEQ history for the full integration solid element (C3D8) and reduced integration solid element (C3D8R). Almost 4x difference in per cycle CEEQ is seen with these two integration schemes. The deformation plot, shown in Figure 13(b), further confirms the wide difference. The out of plane displacement (marked as Y displacement in this figure) of the solder joint at the package corner clearly shows that these two element types give very different results for board deformation. Reduced integration linear solid element is more accurate in capturing the bending deformation. Therefore caution must be used with linear solid element with full integration. However, for smaller packages where bending is insignificant, similar results can be expected for these two integration schemes.
Quadratic 3D hexahedral elements are generally considered most accurate. However, quadratic elements need significantly longer computational times (5~10x compared to the linear elements) and have much larger memory requirements. Linear solid element with reduced integration is a good choice if sufficient element layers are included in the PCB and BGA substrate. Even with more element layers, the linear element models still have much shorter computational time compared to quadratic elements.

Five models of the same BGA package were built to investigate element types and number of element layers. The first model uses 3 layers of linear reduced integration elements in the PCB and 2 layers in the substrate. This model is marked as the “current linear model”. Two other models use 4 and 6 layers of linear reduced integration elements in the BGA substrate and the PCB. The 4th model uses 4 layers of incompatible mode solid elements (C3D8I in ABAQUS) in the PCB and the BGA substrate. Finally, the last model uses quadratic elements with 3 layers in the PCB and 2 layers in the BGA substrate. The averaged per-cycle CEEQ results of these 5 models are shown in Figure 14. Overall, linear element with reduced integration used with 4 or 6 layers gives satisfactory results when compared to the quadratic element.

Incompatible mode linear hexahedral element also provides accurate results, as shown in Figure 14. However, it is very sensitive to element distortions and should not be used if mesh requires irregular shaped elements with small or large interior angles.

Effect of Multiple or Tie Constraint

Mesh continuity between parts of the model is not always easy to achieve. Multi-point constraints (MPC) or Tie constraints serve as convenient tools for mesh transition. However, it is recommended that such constraints be placed away from the location of interest (e.g. solder/pad interface). Figure 15 shows a tie constraint placed above the solder joint between the copper pad and the BGA substrate. This configuration introduces about 40% error in the averaged accumulated creep strain, as shown in Figure 16.

Summary

This paper investigated several aspects in developing an accurate finite element model for obtaining consistent results even if different FEA software packages are used. Volume averaging method is recommended to minimize the stress singularity and mesh sensitivities. Most accurate and
consistent finite element results for the solder joint fatigue in thermal cycling can be achieved when certain guidelines are followed. The following is a summary of our findings and recommendations based on extensive simulation studies:

1. The global/local modeling approach yields satisfactory results if the local model is a cell with one solder ball pitch dimension and no tie constraints are used in the local model other than at the boundaries.

2. Substructure technique can be used without a significant loss of accuracy when most of the structure can be simulated with linear, temperature-independent material properties. However, caution must be used when an assumption of linearization is made.

3. The worst-case solder joint location depends on the selection of the damage metric. It is suggested that the worst-case location be determined by a combination of ‘peel’ stress and averaged creep strain. A proper damage metric (e.g. averaged accumulated creep strain or strain energy density) at the location of interest can then be used to develop the fatigue life prediction model.

4. Regardless of the initial stress-free temperature, the package always adjusts the stress state to achieve stabilized pattern in temperature cycling. Tmax is recommended as the initial stress-free temperature to achieve stabilized solution quickly and to significantly increase the computational efficiency.

5. Linear hexahedral element with reduced integration is recommended if 4 to 6 element layers are included in the PCB and the BGA substrate. Linear hexahedral element with full integration should be avoided, especially if the package and PCB undergo significant bending. Quadratic hexahedral element gives most accurate results but increases the computational time significantly.

Reference


