Classification of location of damage in package-on-package (PoP) assemblies using ANN with feature vectors for progression of accrued damage

CONFERENCE PAPER · JANUARY 2012

DOI: 10.1109/ICPHM.2012.6299535

3 AUTHORS:

Pradeep Lall

Auburn University

460 PUBLICATIONS **2,768** CITATIONS

SEE PROFILE

Prashant Gupta

Intel

28 PUBLICATIONS **112** CITATIONS

SEE PROFILE

Kai Goebel

NASA

340 PUBLICATIONS **3,048** CITATIONS

SEE PROFILE

Classification of Location of Damage in Package-on-Package (PoP) Assemblies using ANN with Feature Vectors for Progression of Accrued Damage

Pradeep Lall, Prashant Gupta
Auburn University
Department of Mechanical Engineering
NSF-CAVE3 Electronics Research Center
Auburn, AL 36849

Kai Goebel
NASA Ames Research Center,
Moffett Field, CA 94035

Abstract— **Miniaturization of electronic products has resulted in proliferation of package-on-package (PoP) architectures in portable electronics. In this study, daisy-chained double-stack PoP components have been used for early-identification of drop-shock impact damage. Time-spectral feature vector based damage pre-cursors have been identified and measured under applied shock stimulus. Experimental strain data has been acquired using strain sensors, digital image correlation. Continuity has been measured suing high-speed instrumentation for identification of failure in the PoP assemblies. The time-evolution of spectral content of the damage pre-cursors has been studied using joint time frequency analysis (JTFA). The Karhunen-Loéve transform (KLT) has been used for feature reduction and de-correlation of the feature vectors for input to an artificial neural network. The artificial neural net has been trained for failure-mode identification using simulated data-sets created from error-seeded models with specific failure modes. The neural net has then been used to identify and classify the failure modes experimentally observed in tested board assemblies. Supervised learning of multilayer neural net in conjunction with parity has been used to identify the hard-separation boundaries between failure mode clusters in the de-correlated feature space. Pre-failure feature space has been classified for different fault modes in PoP assemblies subjected to drop and shock.**

Keywords- Prognostic health management, Karhunen Loéve transform, PoP, neural nets, failure mode classification

I. INTRODUCTION

In this paper, an approach for characterization of damage thresholds and occurrence of dominant failure modes in package-on-package (PoP) components has been presented. A model based-prognostics approach is used for addressing damage initiation and progression along with isolation of dominant failure modes. The framework presents methodologies which are independent of continuous monitoring of daisy chain resistance. The approach is scalable to system level reliability. Currently reactive based fault detection techniques are implements on electronic components such as BIST, fuses, canaries for assessing health of the components. The current trends in industry require low cost smaller, reliable and faster electronic packages. The packaging density as well as more and more of miniaturization of the packages is taking place in electronics industry, hence there is limited space in the X-Y plane of the electronic assembly. 3D packaging with Package-on-Package, Through Silicon-Via (TSV) in the Z-plane is becoming more frequently used. Hence there is need to understand the response/behavior of these assemblies under harsh environment. In this work, drop and shock environment is the harsh environment under which PoP assembly is studies. Though there could be multiple overlapping harsh environments such as thermal cycling, power cycling, drop and shock, aging, it is envisioned that drop/shock environment is the most dominant environment in initiation/onset of damage in critical applications such as avionics and portable electronics. Also with advent of more and more hand help electronics, research community has shifted focus from dominant low cycle ductile fatigue failure of 2nd level solder interconnects under thermal cycling to brittle fracture of 2^{nd} inter-connects as a dominant failure mode in the electronic assemblies under drop. There are no consistent guidelines/trends of failure modes in the current state-of-art for studying assemblies under drop and shock. Hence there is a need for developing and implementing new prognostication techniques for assessing damage levels in electronic-rich systems under shock.

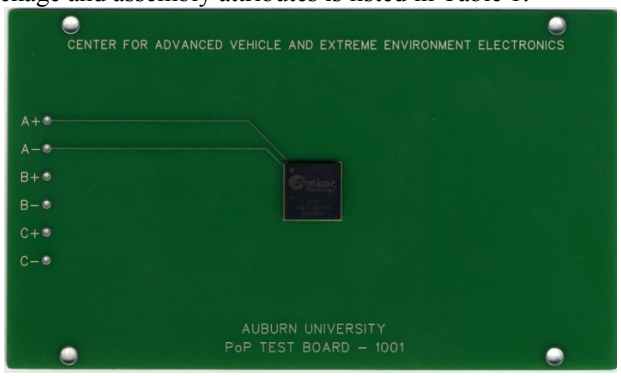
Package-on-package components are typically used for 3D integration of a memory and logic device in handheld electronics. PoP often has a top-tier as a memory device and lower-tier as a logic device. Previous researchers [Yoshida 2006] have outlined stacking guidelines for reliable yield of PoP. [Lin 2006] studied warpage issues in PoP component. Warpage was measured at room temperature and reflow temperatures using Shadow Moire method. Warpage was studied experimentally and by finite element methods for critical package parameters such as die size, substrate thickness, mold thickness and material and design was

optimized. [Cheah 2011] studied novel PoP designs with interconnection between top and lower tier packages using a silicon interposer, package interposer and showed improved electrical performance over conventional PoP design using a solder ball as interconnections between the stacks. [Das 2011] proposed a new breakthrough in PoP design by using Package-interposer-Package (PIP). Das [2011] showed that PIP consisted less warpage in assembly process than a conventional PoP design. This gave flexibility is selection and stacking of various combination of packages for PoP with a PIP approach. [Lin 2011] presented study on PoP as solution to memory application devices in hand held electronics which have high density (up to 16 chips). [Lin 2011] performed finite element simulations on the PoP test vehicle at room temperature and reflow temperature for assessing warpage trends. [Yim 2011] studies warpage of top package in a PoP component with aim of reducing the total height of PoP. [Yim 2011] used compression mold process on various compound and substrate materials to control the height of the PoP in Z direction with minimum warpage at reflow. [Drieza 2006] presented data for board level reliability of Pb-free PoP for thermo mechanical and drop environments. [Drieza 2006] also provided SMT process for defect free yield on PoP.

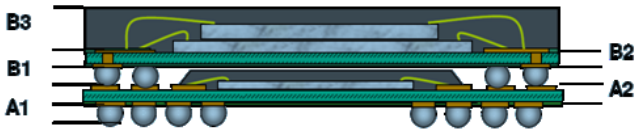
Previously, authors have implemented statistical pattern recognition on feature vectors derived from mahalonobis distance, wavelet packet decomposition, spectral analysis, time-frequency analysis for addressing damage detection and its progression [Lall 2006, 2007, 2008] based on leading indicators of failure in pre-failure feature space. Authors have also developed multi-mode fault isolation techniques using for tracking of dominant failure modes as component level in conventional ball grid array package architectures. Karhunen Loéve transform (KLT) [Lall 2009] and Sammon's mapping [Lall 2011] are used for de-correlation of the feature spaces and parity is presented using supervised [Lall 2011] and unsupervised neural nets [Lall 2010]. In this paper, a conventional PoP design is studied under drop and shock. Solder interconnect is used for connecting top tier memory package with lower tier logic package. Damage thresholds are characterized for initiations and progression of anomalies. Dominant failure modes using neural nets are isolated in pre-failure space populated with feature vectors derived from time-frequency analysis of the transient strain signals obtained from the assembly under shock. Hard parity is achieved between failure modes. Previous researchers have performed research in issues related with micro-fabrication/SMT of PoP, optimizing mechanical and electrical design using interposers, silicon die and solder interconnect, warpage studies at room and reflow temperatures. Prognostics based decision framework for reliability of PoP using leading indicators of failures is new.

II. TEST VEHICLE

A JEDEC format board FR4 assembly for the package-on-package (PoP) has been designed and fabricated at the NSF-CAVE3 Electronics research center surface-mount line (Figure 1). The PoP module consists of two packages including a package-stackable very-fine pitch BGA (PSvfBGA) as the bottom package and a chip-scale package as the top package. Both the top and bottom packages have perimeter array solder interconnect configuration. The bottom PSvfBGA package is a 305 I/O package with a ball pitch of 0.5mm. The top CSP package is a 128 I/O package and has a ball pitch of 0.65mm. Both packages are 12mm x 12mm in size. Height of the assembled PoP package stack is 1.2 mm. A complete set of package and assembly attributes is listed in Table 1.



(a)



Symbol	unit	Min	Nom	Max
A1 (Ball, 0.5 pitch)	mm	0.180	0.230	0.280
A2 (4L laminate)	mm	0.260	0.300	0.340
B1 (Ball, 0.65 pitch)	mm	0.270	0.300	0.330
B2 (2L laminate)	mm	0.180	0.210	0.240
B3 (Mold cap)	mm	0.420	0.450	0.480
Overall Pkg height	mm	1.408	1.490	1.572

(b)

Figure 1: (a) CAVE-3 Package-On-Package Test assembly. (b) PoP Stack with Significant dimensions.

Table 1: Specification/Attributes of PoP Component.

Package On Package	Lower Tier	Top Tier
Package	PSvfBGA	CSP
Solder Alloy	SAC125Ni	SAC105
Size (mm)	12	12
I/O Count	305	128
Pitch(mm)	0.5	0.65
Ball Size (mm)	0.23	0.3
Substrate Thickness (mm)	0.3	0.21
Pad Type	NSMD	NSMD

Each board assembly consists of a single PoP module at the center. Different solder alloys, representative of the actual combinations used in the industry, were used in both top and bottom packages for achieving extended reliability. The alloy composition of the solder interconnects for the top tier package was SAC105 and lower tier package was SAC125Ni. Solder balls between the two tiers of the PoP stack have been divided into two different daisy chain circuits. The first-circuit connecting twelve corner solder balls, three in each corner and

the second-circuit connecting the remaining solder balls. The two daisy chain circuits have been connected to the copper pads on the PCB through two pairs of solder balls from the bottom tier.

III. Development Of Experimental Feature-Vectors

The package-on-package test assemblies were subjected to zero degree JEDEC drop orientation. Figure 2 shows PoP assembly mounted on Lansmont shock system with package-module facing downwards. Both strain-gage and optical digital image correlation were used for measurement of displacement and strain response. Strain gages are mounted at the corners of the package-on-package module on the PoP component.

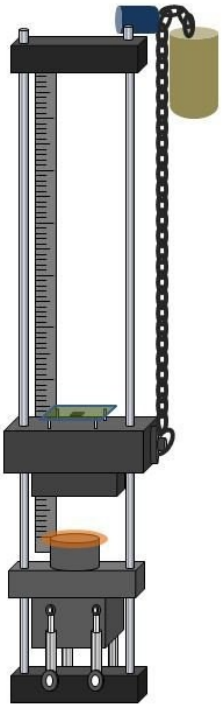


Figure 2: PoP component mounted on Lansmont Shock System in zero degree JEDEC drop orientation.

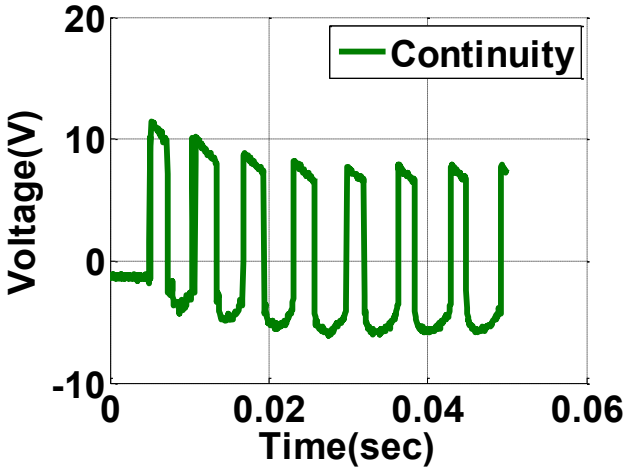


Figure 3: Monitoring of Shock induced damage by Daisy chain resistance.

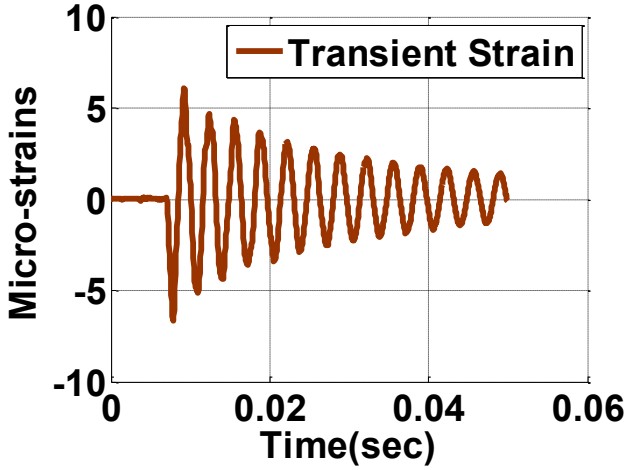


Figure 4: Sample transient dynamic strain from a shock event.

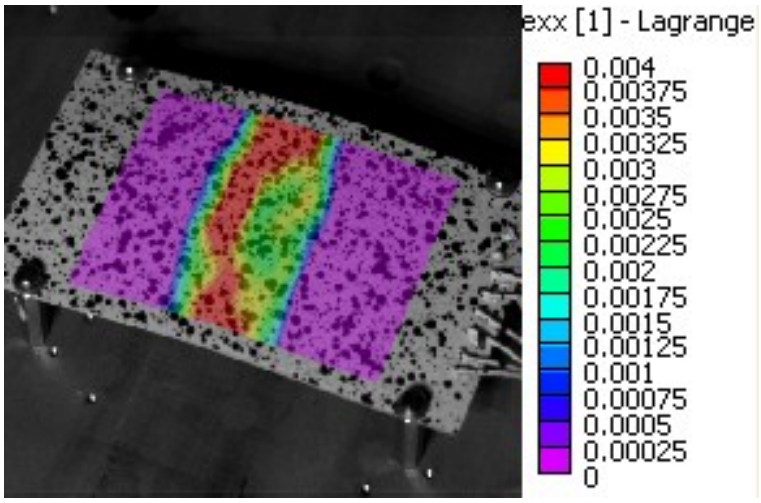


Figure 5: 3D contour of strain in longitudinal direction from DIC of PoP Test Assembly in 0-degree orientation.

Digital image correlation (DIC) was coupled with high speed imaging for full-field measurement of dynamic strain histories under shock event. The board side of the PoP assembly has been speckle coated for DIC. The dynamic response of printed circuit board assembly with PoP component is significantly affected by repeatability of the shock event. Repeatability of the shock event has been quantified using shock-table mounted accelerometer measurements throughout the test. The drop height has been selected such that the assembly experiences a shock pulse of 1500g, 0.5millisecond half sine pulse consistent with the JESD22-B111 Standard specifications. Continuity data has been procured using high speed data acquisition system operating at 2.5 to 5 million samples per second. Figure 4 shows the dynamic strain on the board side of the PoP test assembly. The failure of the PoP component is quantified using increased voltage as indicated by an open daisy chain resistance in Figure 3. High-Speed cameras used for optical measurement have been placed in the vertical plane at an angle of 25 degrees from board-center vertical-axis of the test assembly. The high-speed cameras

have been calibrated before measurement of shock events. A frame-rate of 50,000 frames per second has been used for measurement. Figure 5 represents the sample strain contour acquired from DIC during the drop testing of PoP test assembly. The strain measurements have been used to construct feature vectors for health monitoring from the experimental data.

IV. Development Of Simulation Feature Vectors From Error-Seeded Models

Explicit finite-element models of PoP test assemblies have been developed for JEDEC Drop orientation. The printed circuit board has been modeled with reduced integration shell elements, second-level interconnects have been modeled with Timoshenko-beam elements, and the package elements have been modeled with reduced integration continuum solid elements.

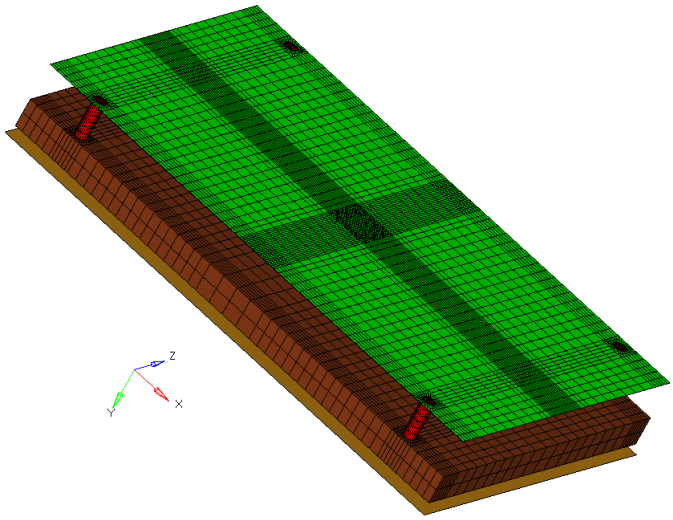


Figure 6: Global Model for Zero Degree JEDEC Drop of PoP assembly.

The board is modeled using reduced integration conventional shell element i.e. S4R. S4R are quadrilateral shell elements which are used for large strain formulation. They have six degrees of freedom, i.e. 3 rotations and 3 translations. The solder interconnects are modeled using two node beam-element i.e. Timoshenko beam elements (B31). These B31 elements have six degrees of freedom, i.e. 3 rotations and 3 translations. To simulate inter-connect behavior; the rotational degrees of freedom are constrained. B31 elements are Timoshenko beam elements and do not preserve the normal behavior of the beam cross section, hence they allow shear deformation of the cross section. This phenomenon of shear deformation is critical in simulation as it is a dominant mode of failure in the first level interconnects. Various layers of the package like copper pad, mold, die, BT substrate have been modeled using C3D8R elements. C3D8R are 8 node reduced integration hexahedral elements with 3 translation degrees of freedom. The floor of impact is modeled using R3D4 element and a weight is attached to the board. The contact between rigid floor and weight of is defined as node to surface using a reference node.

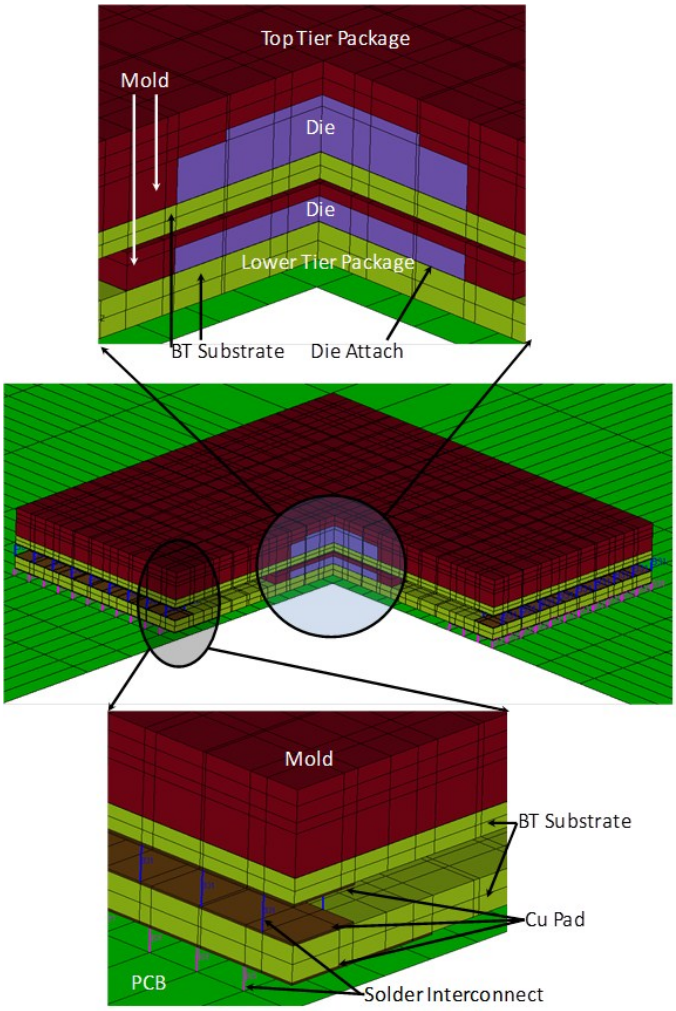


Figure 7: Cross Section of Detailed modeling of PoP module using explicit finite elements on ABAQUS.

The JEDEC orientation for test vehicle A and test vehicle B is shown in Figure 6. Figure 7 shows the cross sectional detail of the modeled packages. Failure modes have been modeled using error-seeding of the finite element model. The modeled failure modes include: solder interconnect failure, solder interconnect partial-fracture, chip delamination and chip fracture take place due to bending during mechanical shock. Strain histories have been extracted from simulation of the shock-event for each failure-mode of the board assembly. The error seeding is performed primarily on the lower tier of the PoP component, since the lower tier component has been found to predominantly experience mechanical shock damage during experiments. Significant strain is taken up by 2nd level interconnects of the lower tier package.

A. Solder Ball failure and Cracking

The solder balls in the drop event are subjected to transient strain due to the considerable bending subjected on the board assembly in a short period of time. The solder balls undergo

fatigue failure leading to complete interconnect-failure or partial solder-interconnect damage under repeated shock events. Previous researchers have shown that the corner most solder balls in the package will experience maximum shear-strain. Most commercial packages use one or more solder interconnects as redundant ground pins. Corner redundant ground pins facilitate damage monitoring based on continuity-measurement of the corner solder balls. Since the corner interconnects have the maximum distance from the neutral point, hence they are most likely to fail in the drop events.

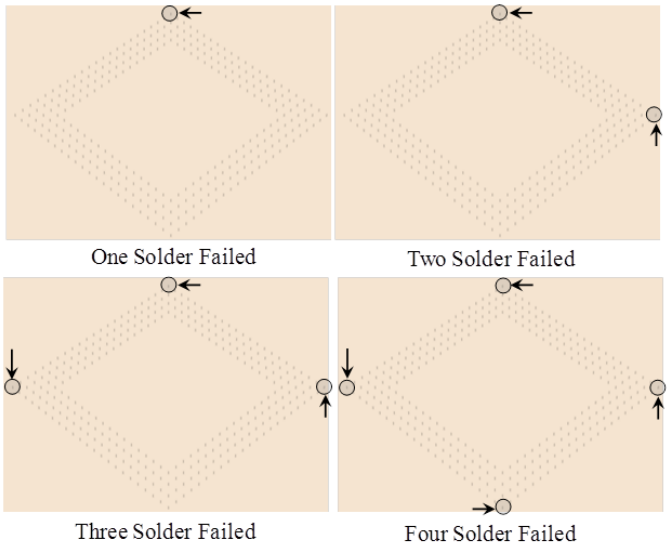


Figure 8: Error seeded Explicit Finite Element Models: Solder Interconnect Failed.

Solder ball crack initiation and propagation and reduction in the effective cross-sectional area has been simulated on lower tier of the PoP component using explicit model by reducing the cross sectional area of the modeled solder interconnect. Simulation of complete failure of the solder ball has been simulated by elimination of the solder interconnect from the model. Figure 8 shows a schematic of solder interconnect array of the lower tier package simulated with successive failure of additional corner solder balls in PoP test vehicle. Inter-connect failure is modeling by sequentially selecting solder balls at the periphery of the inter-connect array configuration. A comprehensive simulation data set is formed by simulation of inter-connects failure. Transient strain histories have been obtained for the drop events from elements at the center and corners of the package (on the mold side as well as on PCB side).

B. Die Cracking

Chip fracture has been simulated on PoP component in the lower tier package. Die cracking has been modeled at the underside of the chip. The fracture on the chip is modeled as a crack occurring in the chip. The crack has been modeled by removing elements at the crack location and replacing them by a contact surface. The contact surfaces represent in the die crack. The chip fracture mode of failure is shown in Figure 9.

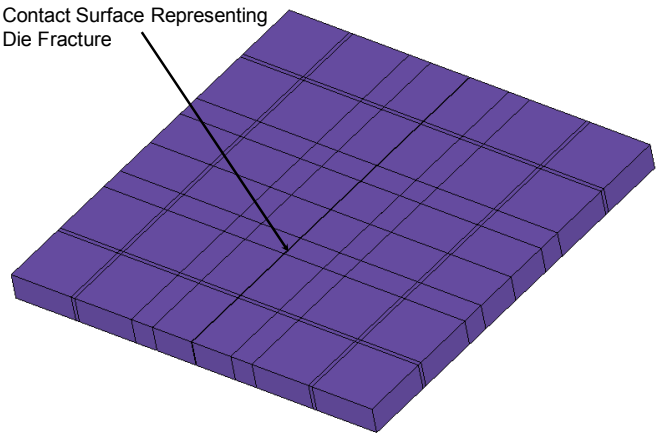


Figure 9: Error seeded Explicit Finite Element Models: Die Cracking.

Chip Delamination

Chip delamination is also a very common mode of failure occurring in the drop events. Chip Delamination occurs when there is a crack or detachment between the bond in the chip and the BT substrate. The phenomenon is modeled by inclusion of contact elements between the die attach and bottom-package BT substrate of the PoP component. Figure 10 shows the model of chip delamination in the PoP component.

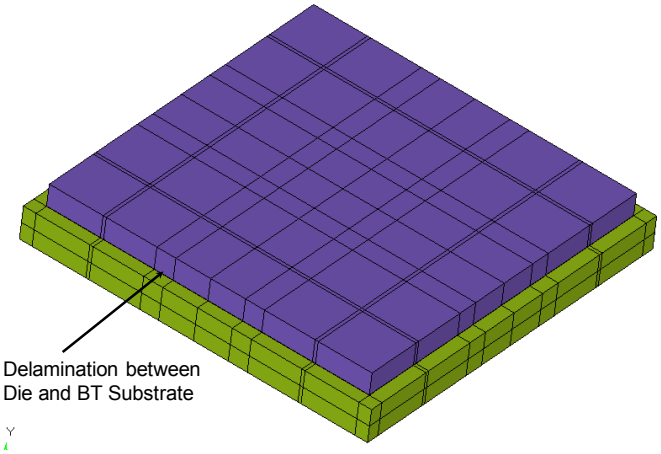


Figure 10: Error seeded Explicit Finite Element Models: Chip Delmaination.

C. Part Fall Off

Catastrophic failure during mechanical shock may result in part fall-off from the printed circuit board assembly. Part fall off is also simulated as one of the possible failure modes in this study. While the part fall-off will result in loss of functionality, this failure mode has been included in the study as a limiting case. The change in the feature vectors has been studied for various system state including healthy modules, incremental damage and catastrophic failure. Strain histories at the board side near the center and corner of the package-location have been extracted. Figure 11 shows the

configurations of the test boards used in this research work with PoP part falling off the PCB assembly.

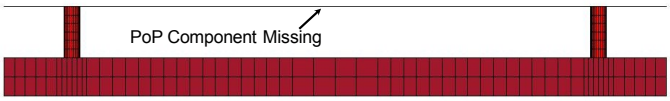


Figure 11: Error seeded Explicit Finite Element Models: Part Fall off.

V. Mapping Failure Modes In De-Correlated Feature Space

The methodology developed in the study for failure mode identification in PoP electronic assembly under mechanical shock can be used to classify multiple failure modes in the pre-failure space. Karhunen-Loéve Transforms (KLT) is used for de-correlation of the feature space, which isolates the center of gravity of each impending failure mode in the feature space. Multilayer Perceptron feed forward neural network has been designed to extract hard decision boundaries between competing failure-modes in the feature space. Figure 12 represents the schematic of various data processing sequences for performing real time fault monitoring in PoP test assembly under transient drop and shock loads. In this study, a feature vector based approach has been adopted. Continuous stream of time domain data is used as the in input for feature extraction. Time domain data is used for feature extraction using joint time frequency analysis (JTFA) of time domain dynamic transient strain for extraction of time-moments and frequency-moments (Figure 12). The JTFA feature space has been de-correlated using Karhunen Loéve Transforms (KLT). Hard separation and identification of the region in the feature space belonging to each failure mode is extracted using a supervised multi-layer perceptron neural net.

Various dominant failure modes in the PoP test vehicle subjected to transient drop have been isolated in this work. Two comprehensive data sets have been procured for the PoP test vehicle. The error seeded simulations of the test vehicles have been performed for various targeted failure modes such as pristine state (healthy), solder inter-connect cracking, solder inter-connect completely failed, chip cracking, chip delamination, and part fall-off. The error seeded simulation data set is obtained by extracting strain histories from PoP test assemblies which have been error seeded with known magnitudes of damage. Strain histories from board side as well as from package sides have been captured at the location of PoP component-corners on the test-assembly. Only one failure mode has been targeted in each simulation. Experimental data set has been procured by performing shock testing of the PoP test assemblies in JEDEC configuration. The strain response is captured at various target locations on the test board. Both contact (strain gage data) as well as non-contact (high speed imaging in conjunction with digital image correlation) measurement of the strain histories is performed. The assembly is tested from pristine state to some accrued state of damage as well as beyond failure in experimentation.

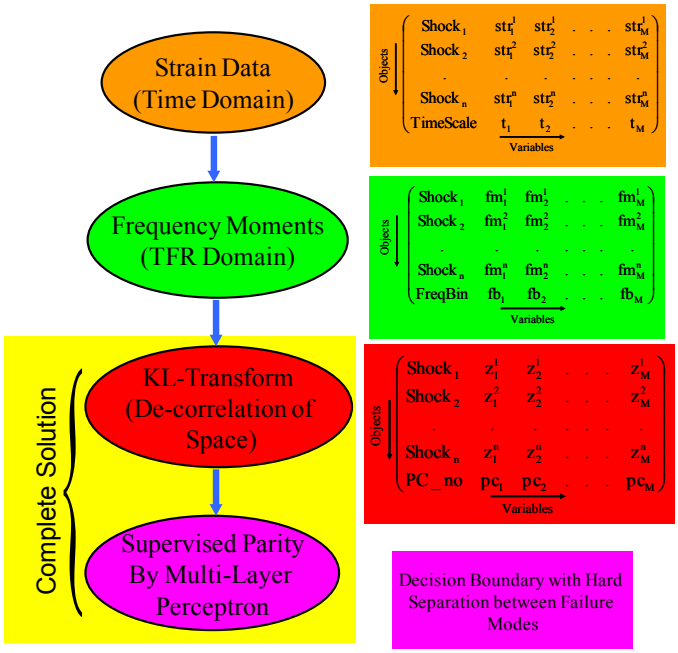


Figure 12: Data Processing Sequence for Real Time Fault Monitoring.

A. *Joint Time Frequency Analysis for Shock Signals*

Evolution of the frequency content in the transient response of the PoP circuit board assembly has been analyzed using joint time-frequency analysis. Previously, the authors have implemented time-frequency feature vectors as a damage proxy for quantification of damage initiation and progression using statistical pattern recognition [Lall 2007, 2008] in test assemblies under transient drop. Since the transient strain history response from the test assemblies are in time domain and non-stationary in nature, hence little can be inferred how frequency content is evolving with time in the signal. To exploit relevant variance in the transient strain signal, JTFA has been used. Cohen class of transforms has been applied to compute the joint time frequency distribution. A reduced interference distribution (RID) kernel has been used as an auxiliary function to reduce cross-terms and thereby reducing the interference, which is seen in other popular joint time frequency analysis techniques such as Wigner-Ville transforms [Cohen 1989, 1995, Williams 1994]. In this study, the binomial time-frequency kernel proposed by [Jeong 1992[a,b]] has been applied to study the drop and shock characteristics of an electronic assembly. The binomial time-frequency Distribution defined by [Jeong 1992[a,b]] is,

$$\mathrm{TFR}(n,\nu) = \sum_{\tau=-\infty}^{\tau=\infty} h(\tau) \sum_{\nu=-|\tau|}^{\nu=+|\tau|} \frac{g(\nu)}{2^{2|\tau|}} \binom{2|\tau|}{|\tau|+\nu} f(n+\nu+\tau) f^{*}(n+\nu-\tau) e^{-i4\pi\omega\tau} \quad (1)$$

where $h(\tau)$ and $g(\nu)$ is the time smoothing window and the frequency smoothing window respectively and $f(n)$ represents the signal where $n = 1,2 \cdots N$. The term $\nu = \theta\tau$, and is used to define the RID kernel as the RID kernel

constraint is that $\theta\tau >> 0$. The frequency smoothing window $g(\nu)$ and the time smoothing window $h(\tau)$ used here is a hamming window of size (N) as outlined in [Jeong 1991, Jeong 1992[a,b], Williams 1994]. Figure 13 shows a time frequency distribution of a transient strain history obtained from the JEDEC drop of the test vehicle.

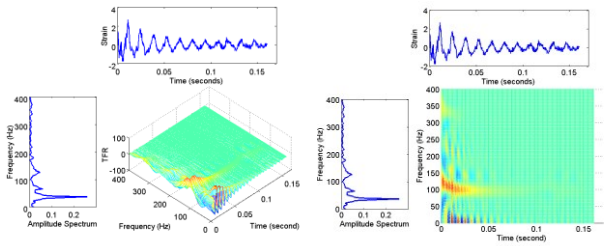


Figure 13: Joint time Frequency Feature Space of a Transient Strain signal.

The time frequency analysis of the signal has been used to obtain the frequency content of the transient strain signal at each given time instant. The time frequency distribution obtained for a JEDEC standard horizontal drop of an electronic assembly is shown in Figure 13. The time-frequency signature is based on transient strain signal obtained from two sources including the strain sensors placed on the electronic assembly and digital image correlation based on high-speed imaging during the shock event.

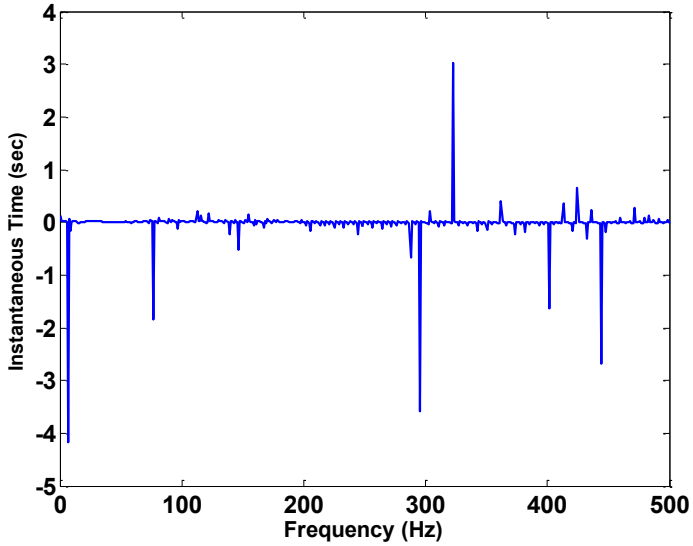


Figure 14: (a) Sample Frequency Moment Feature Vector for Transient Strain signal from a JEDEC drop

The time moment and frequency moment distributions shown are unique to a given signal and represent the strain signals in the joint time-frequency spectrum. The first order moments, in time and in frequency, of a time-frequency energy distribution, tfr, describe the averaged positions and spreads in time and in frequency of the signal. The time moment represents an estimation of instantaneous frequency at a given time instant during the drop event [Boashash 1989, Cohen 1995, Tacer 1995]. The frequency moment represents an estimation of the group delay of the signal for each frequency in the signal. [Cohen 1995, Tacer 1996, Georgopoulos 1997]. Frequency moments are proportional to time-frequency distribution (equation (2)); hence they are used as representative feature to study evolution of the variance in the feature space with shock events in time. Since frequency moment represented by equation (3), feature vectors are unique for each signal they are an appropriate choice for prognostics of electronic assemblies in drop and shock. Figure 14 shows the moment feature vectors obtained for the transient strain signal respectively

First Frequency-moment

$$t_m(f) \propto tfr(t,f) \quad (2)$$

$$t_m(f) = \frac{\int_{-\infty}^{\infty} t\ tfr(t,f)\,dt}{\int_{-\infty}^{\infty} tfr(t,f)\,dt} \quad (3)$$

B. De-Correlation of the Feature Space by Karhunen Loéve Transforms (KLT)

KLT is a statistical classifier which has been used for de-correlation of feature space extracted from JTFA analysis of the transient shock signals. KLT is a linear projection scheme which leads to abstraction of information and dimensional reduction. This property of KL transform is exploited in this work for studying damage progression in package interconnects during successive drop events. The de-classification of feature space has been done based on the variability of the data. The time-frequency feature space has been clustered into most dominant directions of variability. Previously, KLT has been used for data compression in classical communication theory [Ogawa 1986, Yamashita 1992, Shawn 2004]. The use of KLT for failure-mode classification of 3D packaging components such as Package on Package is new. The data set described in Figure 15 has been de-correlated using KL transform.

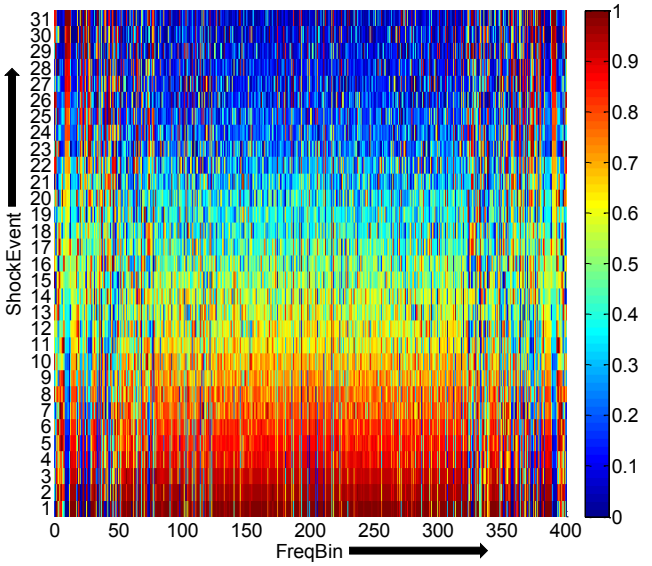


Figure 15: Sample Classification Matrix for fault isolation.

Let X is the representation of the variable-space in the environment of interest. Joint Time frequency distributions of the strain histories from successive drop events from pristine assembly configuration to failure for each board assembly has been used as the input matrix, X. A de-classified feature space Z has been obtained using the KL transform of the matrix, X. The de-classified set of vectors is a linear combination of principal components with decreasing order of importance. The initial k-vectors are important as they account for most of the variance in the data.

$$[X] = \begin{pmatrix} \text{Shock}_1 & \text{fm}_1^1 & \text{fm}_2^1 & . & . & . & \text{fm}_M^1 \\ \text{Shock}_2 & \text{fm}_1^2 & \text{fm}_2^2 & . & . & . & \text{fm}_M^2 \\ . & . & . & . & . & . & . \\ \text{Shock}_n & \text{fm}_1^n & \text{fm}_2^n & . & . & . & \text{fm}_M^n \\ \text{FreqBin} & \text{fb}_1 & \text{fb}_2 & . & . & . & \text{fb}_M \end{pmatrix} \quad (4)$$

$$[X]_{m-\text{dimensions}} \overset{KLT}{\rightarrow} [Z]_{m-\text{dimensions}} \quad (5)$$

The data set X has been centered and scaled to eliminate a non-zero mean of matrix X. 'Since the input matrix has been centered and scaled, the expected value of the input matrix is zero,

$$E[X] = 0 \quad (6)$$

The input matrix X has been projected on a unit vector, q in the de-correlated feature space, also of dimension-m. The projection is represented as an inner product of vectors X and q is the matrix of principal components, A,

$$A = X^T q = q^T X \quad (7)$$

The variance of A has been represented as a function of unit vector, q,

$$\begin{aligned} \sigma^2 &= E[A^2] \quad (8) \\ &= E[(q^T X)(X^T q)] \\ &= q^T E[XX^T] q \\ &= q^T R q \end{aligned}$$

Where, the outer-product of X, has been represented as R. The variance probe, Ψ has been computed, since it relates correlation matrix, R, the unit vector in de-correlated space, q, and standard deviation, σ,

$$\psi(q) = \sigma^2 = q^T R q \quad (9)$$

The equation that governs the unit vectors, q, and variance probe Ψ an eigenvalue problem, which has a non-trivial solution, $q \neq 0$,

$$Rq = \lambda q \quad (10)$$

The eigenvalues have been arranged in descending order,

$$\lambda_1 > \lambda_2 > \lambda_3 > \ldots\ldots > \lambda_m \quad (11)$$

The eigenvector matrix Q is orthogonal consisting of column-vectors which satisfy the condition of orthonormality,

$$Q = \begin{bmatrix} q_1 & q_2 & q_3 & \ldots & q_n \end{bmatrix} \quad (12)$$

$$q_i^T q_j = \begin{cases} 1 & j = i \\ 0 & j \neq i \end{cases} \quad (13)$$

Using eigen-decomposition, the correlation matrix R has been written in terms of eigenvalues and eigenvectors as,

$$R = \sum_{i=1}^{m} \lambda_i q_i q_i^T \quad (14)$$

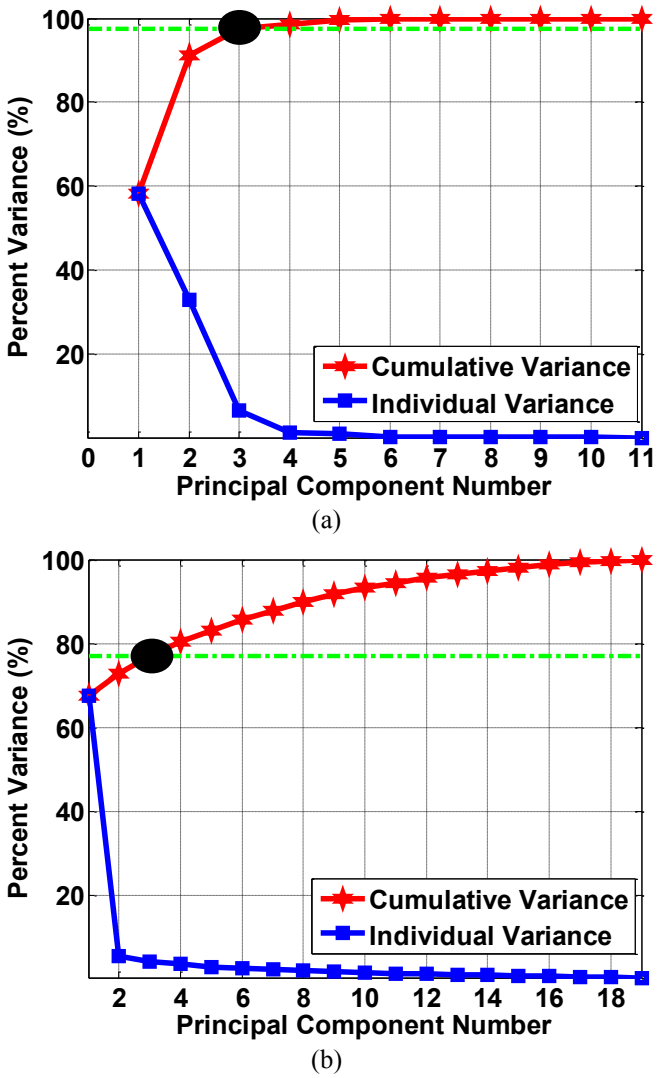


Figure 16: Scree Plots of PoP Test Board (a) Simulation (b) Experiment.

The eigenvectors of the correlation matrix R represent the principal directions along which the variance probes $\Psi(q_j)$ have extreme values. The associated eigenvalues define the extreme values of the variance probe $\Psi(q_j)$. There are m possible projections of x, corresponding to m possible solutions of unit vectors q. Projections a_j , which are the principal components, have been combined into a single vector as follows,

$$A = [a_1, a_2, a_3, \ldots\ldots\ldots, a_m]^T \quad (15)$$

$$A = [x^T q_1, x^T q_2, x^T q_3, \ldots\ldots\ldots, x^T q_m]^T \quad (16)$$

$$A = Q^T x$$

The original data-vector has been synthesized from the transformed feature space of principal components by pre-multiplying the above equation by Q,

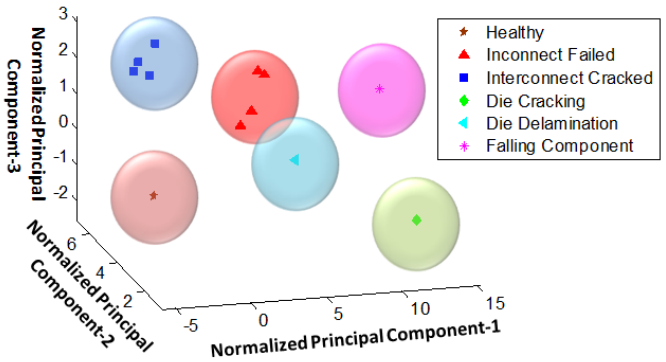
$$X = QA \quad (17)$$

$$X = \sum_{j=1}^{m} a_j q_j$$

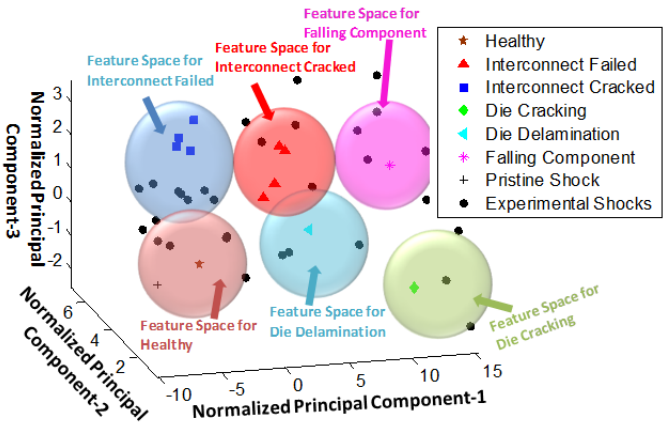
The original data vector "x" has reduced dimensions from the transformed feature space "a" as,

$$\hat{x} = \sum_{j=1}^{\ell} a_j q_j \qquad \hat{x} = \begin{bmatrix} q_1 & q_2 & . & . & q_\ell \end{bmatrix} \begin{bmatrix} a_1 \\ a_1 \\ . \\ . \\ a_\ell \end{bmatrix} \quad (18)$$

where $l < m$. The KL transform has been used to create a linear projection of feature space from m- dimensions to l – dimensions, which approximates the original data x. Dominant directions of the de-correlated feature space have been determined by first few largest eigenvalues.



(a)



(b)

Figure 17: De-correlated Space (a) Error Seeded Data Set (b) Experimental Data set overlapped on Error seeded data.

The principal components entering transformation are also determined by the dominant eigenvalues. Figure 16 represents scree plots of PoP Test board for simulation and experimental data sets respectively. Figure 17a represents the de-correlated errors seeded data set populated in first three dimensions. Distinct groups in various fault modes can be seen in separation in Figure 17a. Figure 17b represents the de-correlated overlap plot of experimental feature space on error seeded space. Various clusters of different dominant failure modes get populated with experimental features Figure 17b. In this study, the idea is to develop a prognostics based methodology for fault monitoring as well as damage diagnostics. Development of supervised neural controller is performed next in this work.

VI. Mutlilayer Perceptron Network For Hard Parity Of Failure Modes

Stepwise parity of failure modes has been performed on a 3-dimensional de-correlated space obtained by KL transform on the PoP test assembly. A multilayer perception network has been developed using supervised training of perceptrons. Perceptrons have been trained using delta learning rule. Since simulation data set consists of shock features from drop simulations exclusively error seeded with a certain type of failure. This exclusive error seeding of failure modes facilitates the identification of the failure mode location in the feature space. Each fault mode is taken as a potential target in-turn and segregated from other potential fault modes in the 3D feature space. The training of perceptrons using delta rule is as explained as follows.

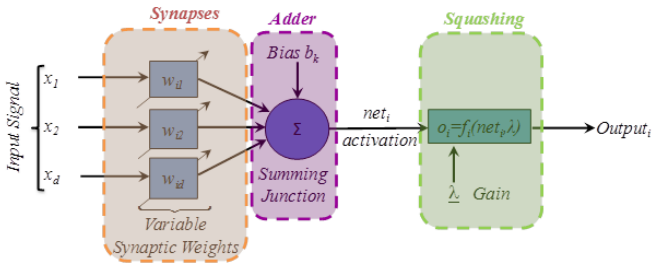


Figure 18: Generic Perceptron Layout.

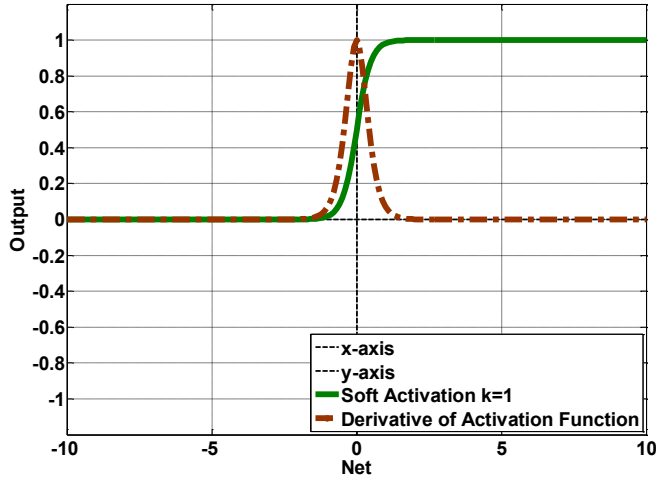


Figure 19: Soft Sigmoidal Activation Fuction for a Unipolar Neuron.

A. Delta Rule

A single neuron can separate 2-classes with a line or a plane or from a hyperplane depending on the dimensionality of the input space. In this work, perceptrons are used for separating different dominant failure modes using a plane in a 3D feature space. Training of a perceptron with delta rule generalizes to back propagation for a multi-layer perceptron feed-forward network. Figure 18 shows a generic single perceptron used for parity of classes (fault modes) in this study. The training scheme has 3 main phases i.e. initialization, learning and adaptation. Each phase is explained as follows.

B. Initialization

In this study, sequential training of perceptrons is performed for stepwise parity. Each pattern in the error seeded data set fires randomly in the neuron. Before training starts various parameters are initialized. The weights of the input synapses have been initialized randomly. Learning rate (η), gain across the neuron (k), maximum error (e_{max}), number of iterations (n_{max}) are initialized. A soft-activation function has been selected. In this study, unipolar neurons are selected with a soft sigmoidal activation function as shown in Figure 19. The output of unipolar neurons varies between 0 and 1 as seen in Figure 19. Equation (19) represents the soft sigmoidal activation function, and Equation (20) represents calculated by the perceptron using the soft activation acting on the input vector [Bishop 1995, Haykin 1999, Rumelhart 1986, Schalkoff 1992]:

$$\varphi(v) = \frac{2}{1+e^{-2kv}} - 1 \tag{19}$$

$$y_j = \varphi(v_j(n)) \tag{20}$$

Where, k is the gain of the neuron, v is the induced field, and y is the net from the summing junction. When patterns are passed to the perceptron, adder operation is performed, where initialized weights of the synapses are linearly combined with the input pattern. The net output of the single layer perceptron unit when presented with a training pattern is given by Equation (21), which is inner product between the input pattern and the weight vector [Bishop 1995, Haykin 1999, Rumelhart 1986, Schalkoff 1992]:

$$v_j = \text{net} = \sum_{i=1}^{m} w_{ji}x_i + b_j = \sum_{i=0}^{m} w_{ji}x_i \tag{21}$$

$$\text{where, } x_0 = +1 \text{ and } w_{j0} = b_j$$

Since delta learning rule is a supervised training scheme hence before training is performed a target is also initialized. Here since stepwise parity of each failure mode is performed in the feature space, hence each failure mode is taken as a potential target in the design of a multilayer perceptron network.

C. Learning

Learning is one of the primary stages of Delta rule. The net across the layer of the neuron is passed through the activation function (soft sigmoidal in this study) for output. Activation function limits the amplitude of the output of a perceptron. This limiting property of the output of the perceptron is called squashing which keeps the output value of the perceptron within permissible limits (0-1: for unipolar neurons).

$$y_p = \varphi(w_1x_{p1} + w_2x_{p2} + ... + w_nx_n) = \varphi(\text{net})$$

Back propagation algorithm for training of single unit in this study is based on optimization of a criterion function by iterative gradient descent method. The criterion which is used to terminate training of the perceptron unit, is defined as the sum of squared error (SSE) or total error energy (ξ). The criterion function is given by Equation (22), where e_p is the output error vector between the desired response from the perceptron and the measured response, and is given by Equation (23) [Bishop 1995, Haykin 1999, Rumelhart 1986, Schalkoff 1992]:

$$\xi(n) = \frac{1}{2}\sum_{j\in C} e_j^2(n) \tag{22}$$

$$e_j(n) = d(n) - y(n) \tag{23}$$

Where total error energy is calculated for the j^{th} neuron for the n^{th} iteration. Hence total error energy function is proportional to synaptic weights as well as bias levels of the perceptron. The learning of the perceptron unit will optimize total error function in such a way that it is minimized.

D. Adaptation

In this part of the training the synaptic weights are updated such that the objective function as given in Equation (22) is optimized. The amount of correction applied to the weight vector is proportional to the derivative of the total energy w.r.t. the weight vector [Bishop 1995, Haykin 1999, Rumelhart 1986, Schalkoff 1992]:

$$\Delta w_{ji}(n) = -\eta\frac{\partial\xi(n)}{\partial w_{ji}(n)} \tag{24}$$

$$\Delta w_{ji}(n) = -\eta\frac{\partial\xi(n)}{\partial e_j(n)}\frac{\partial e_j(n)}{\partial y_j(n)}\frac{\partial y_j(n)}{\partial v_j(n)}\frac{\partial v_j(n)}{\partial w_{ji}(n)} \tag{25}$$

$$= -\eta.e_j(n).(-1).\varphi_j'[v_j(n)].y_i(n)$$

Where η is the learning rate of the training by back propagation. The negative sign represents direction of weight change such that the total error energy function $\xi(n)$ minimizes. The final weight update equation is given by Equation (26):

$$w = w + \Delta w \tag{26}$$

Each targeted failure mode is partitioned in the feature space and represented by a single layer perceptron unit. The pseudo-code shown below represents the flow of the training algorithm used for supervised parity of the feature space for each targeted dominant failure modes.

VII. Hard Parity Of Fault-Modes In Pop

The targeted failure modes in this study are the pristine stage, stages with solder interconnect cracking, solder interconnect completely failed, chip delamination, die cracking, part fall off. It is seen from previous experience of the authors [Lall 2008, 2009, 2010, 2011], that above failure modes are

representative of a global set of failure modes, which cover most of the damage scenarios (partial, complete as well as mixed modes) in the electronics subjected to drop and shock. Table 2 shows the truth table used for stepwise parity of die delamination.

Table 2: Truth Table Used for Chip-Delamination

Failure Mode	Label	Desired Output
Healthy	P	0
Solder Cracking	IC	0
Solder Missing	IM	0
Die Cracking	***DC***	0
Die Delamination	DD	***1 (Target)***
Fall Component	FC	0

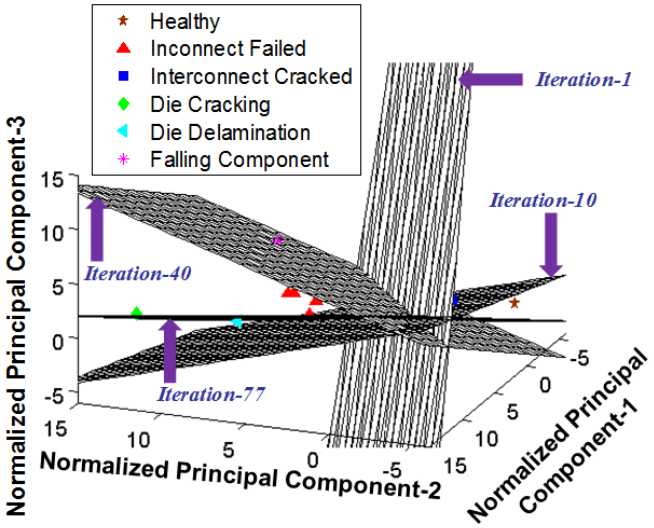
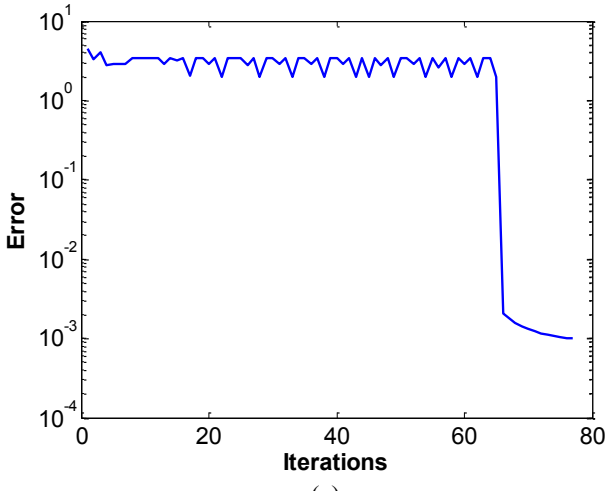


Figure 20: Parity by Planes in 3D space for die delamination.



(c)

Figure 21: Error Plot for training of perceptron representing die delamination.

Similar truth tables have been used for other failure modes. Figure 20 represents the training of the single layer perceptron used for parity of the targeted failure modes under consideration i.e. die delamination. Before training is performed the weights of the perceptron are initialized randomly. These weights represent an equation of plane as parity is performed in the first 3 dimensions of the principal components representing the pre-failure feature space. Figure 20 shows that as training proceeds, the weights of the neurons evolve represented by evolving of the plane. The training stops when an optimum set of weights (plane) is achieved as the objective function is optimized. Figure 21 shows the error plots of the training across single layer perceptron for targeted failure modes of the PoP test assemblies. Once stepwise parity of the feature space has been performed, decision boundaries between different regions of the feature space representing different failure modes can be readily formed. In this study, since first 3 principal components are used for populating the feature space hence, the boundaries representing parity between dominant failure modes are given by a plane. The regions belonging to each failure mode in the feature space are volumes. A hyper-region (volume) is extracted from the feature space by performing the AND operation [Jain 1996, Wilamowski 2007] on the layer of neuron used for stepwise parity of the feature space. Figure 22 shows the schematic of the artificial neural network designed for identification of the region belonging to die delamination in the feature space. The designed network is a 3-5-6-6 networks i.e. (3 inputs, 5 neurons in 1st hidden layer, 6 neurons in 2nd hidden layer and 6 outputs). The weights of the neurons representing the hyper region are tuned using simulated annealing.

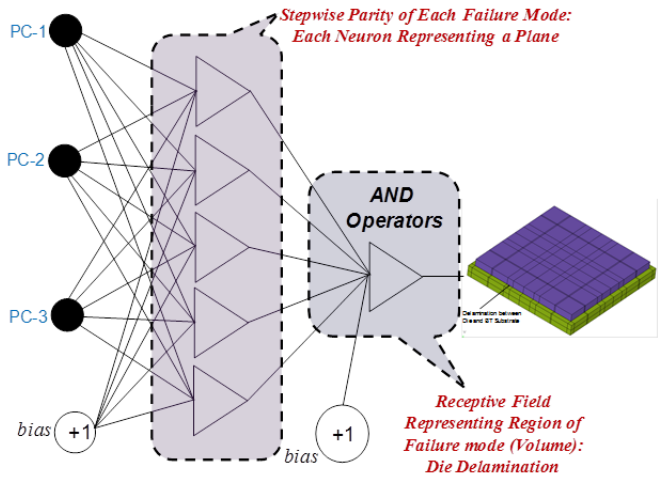


Figure 22: Schematic of Designed Neural Network used for Isolation of a single Fault Mode (Die Delamination) in PoP test assembly.

VIII. Multiple Failure Modes In Feature Space

In this section, the process for inference of the classified failure-modes is explained for the experimental data set and error seeded simulation data set. Experimental data set is obtained by subjecting PoP test-vehicle to successive shock events. The data set captures damage features from pristine state to a partially failed state, and eventually to a completely failed PoP component.

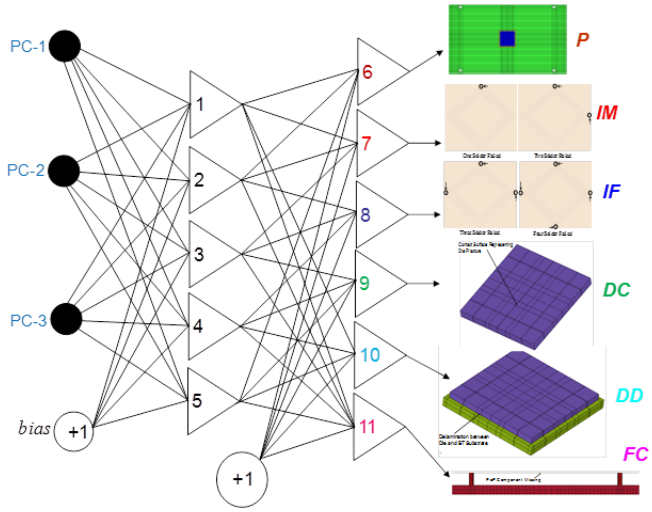


Figure 23: Final Designed Neural Network for Health Monitoring of PoP Test Assembly.

Table 3: Labels for Different fault scenarios.

Label	Anomaly
P	Healthy
IM	Interconnect Missing
IF	Interconnect Fracture
DC	Die Cracking
DD	Die Delamination
FC	Falling Component

The data sets are multidimensional in nature. Error seeded simulation data set is obtained by simulating JEDEC drop using explicit finite element modeling of the PoP test-vehicle. Various failure modes are exclusively error seeded in each simulation. The simulation data set consists of features from pristine assembly, partial and complete failure of 2nd level solder interconnects along with failures inside the PoP component from die cracking and chip delamination. Since drop and shock of electronic components are low cycle fatigue phenomenon hence dynamic strain is chosen as a proxy for quantifying damage initiation and progression along with fault isolation. Since each failure mode is exclusively error seeded in each finite-element simulation, it is known ahead of time which shock event in simulation classification matrix belongs to what kind of failure.

Table 3 shows the labels for various error seeded failure modes taken under consideration in this study. Figure 23 represents the supervised neural network of perceptrons developed in this work. It is a 3-5-6-6 network i.e. with 3 inputs, 5 hidden neurons in layer-1 and 6 hidden neurons in layer-2 and 6 outputs. The three inputs signify the first three principal components extracted from time-frequency analysis in conjunction with KL transform. The 5 hidden neurons in layer-2 represent equations of stepwise parity of each targeted failure mode using planes in a 3-dimensional pre-failure space. The 6 hidden neurons in layer-2 represent region in the 3-dimensional feature space (volume) belonging to each failure mode. The output combinations of 6 neurons are used for fault isolation. Figure 24 shows the classified fault modes in the error seeded data set populated in first 3 principal directions. Various targeted failure modes are seen in different regions of the feature space. Since parity is performed in 3 dimensional space hence the perspective region of volume belonging to each failure mode is seen in different rotations of the feature space as demonstrated from Figure 24a-f.

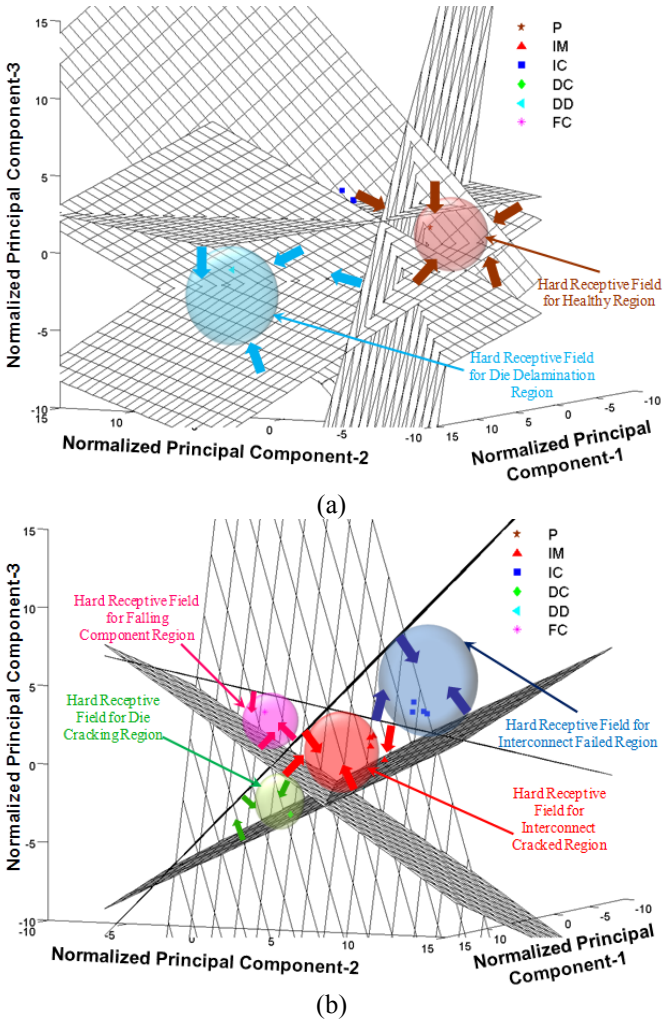
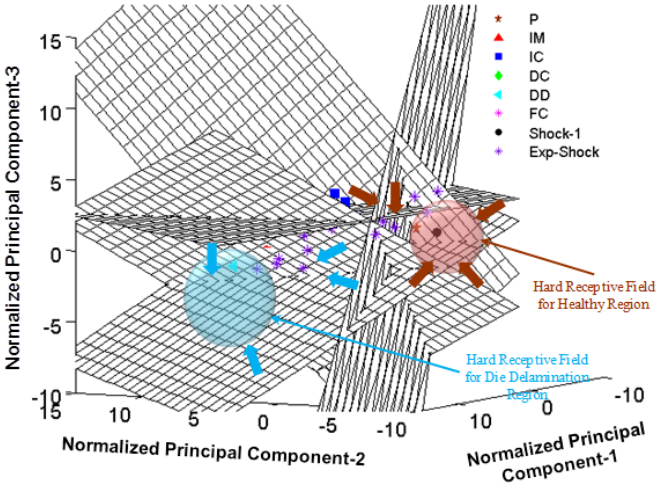
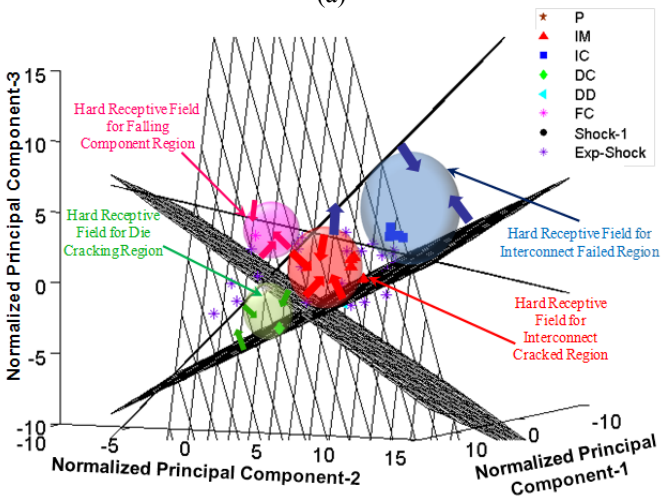


Figure 24: Harp Parity of Error seeded feature space populated in 3 dimensions for various failure modes (a) View-1: Represents region of feature space belonging to pristine state and die delamination, (b) View-2: Represents regions of various fault mode (eg. Interconnect cracking, interconnect missing, die cracking, falling component).

The classification seen here is based on severity of each failure mode as error seeded simulations are exclusively error seeded. In a practical scenario i.e. in experimental data set multiply failure modes can occur simultaneously. Also with each shock event in time the severity of each failure mode can change with accrued damage. Figure 25 represents the overlapped experimental feature space on simulation space. It is known ahead of time which region in failure space belongs to which failure mode.



(a)



(b)

Figure 25: Overlap of error seeded space and experimental space populated in 3 dimensions. (a) View-1: Represents region of feature space belonging to pristine state with pristine experimental shocks and die delamination, (b) View-2: Represents regions of various fault mode (eg. Interconnect cracking, interconnect missing, die cracking, falling component).

Hence various shock events populate the first 3 principal directions of the failure space. It is seen that volume of the failure space belonging to solder interconnect cracking and failed is maximum. This is expected because of the dominance of the second-level interconnect damage in PoP assemblies when exposed to mechanical shock. There are some regions which do not consist of any shock feature or where no simulation feature exists. These regions are regions where some unexplored failure-modes may exist. Isolation of different failure modes is based on severity hence any shock event lying in a particular region of the space signifies that a particular failure mode is initiated and propagating with every shock event in time. Hence the approach presented here is inclusive of tracking damage progression and isolation.

IX. Validation Of Classification

In this section, the results of classification of different failure modes are validated. The classified failure modes are validated using statistical hypothesis test and experimental cross sections of the failed PoP component.

A. Multivariate Analysis of Variance

In this study, Multivariate analysis of variance is used to quantify statistically significance of classification. The experimental data sets and simulation data sets used in this work are high dimensional data sets with multivariate nature. MANOVA is used for performing difference in means of different groups/classes (fault modes) in a 3 dimensional matrix form. Mean being a measure of central tendency of the data, hence if the means of different inherent classes is statistically not different then it signifies that the classes are completely diluted in the data-set, hence classification of the different patterns in the data will yield no meaningful result. MANOVA is performed to check if the means of different classified failure modes in a 3-dimensional feature space are statistically different or not. Wilk's lamda is the test statistic used for hypothesis testing. Likelihood ratio test is used for finding Wilk's Lamda test statistic. It is given by,

$$\Lambda = \frac{\text{ConstrainedML}}{\text{UnconstrainedML}} = \left(\frac{\left|\hat{\Sigma}\right|}{\left|\hat{\Sigma}_0\right|} \right)^{n/2} \tag{27}$$

Where, $\hat{\Sigma}$ is the estimate of covariance by Maximum Likelihood Function under alternate hypothesis and $\hat{\Sigma}_0$ is estimate of covariance by Maximum Likelihood Function under null hypothesis. Likelihood Function is given by:

$$L(\mu, \Sigma) = \frac{1}{(2\pi)^{np/2}\left|\Sigma\right|^{n/2}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} (x_i - \mu)' \Sigma^{-1} (x_i - \mu)' \right] \tag{28}$$

Where, n is the degree of freedom, p is the dimensionality, Σ is the estimate of covariance and μ is the mean.

Table 4: Significance of classification by Multivariate analysis of variance (MANOVA).

	Statistic	Value	F-Value	Pr>F
Error Seeded Model	Wilk's Lamda	0.00005	26.11	<0.0001
Experiment	Wilk's Lamda	0.0609	5.62	<0.0001

Table 4 represents the results of MANOVA performed on the features which populate the partitioned 3-dimensional pre-failure space with different regions of the feature space belonging to different fault modes. A p-value less than 0.05 signifies that statistically there is 95% confidence at least one of the means of the classified failure mode is significantly different from the means of the other classified failure modes.

B. *Similarity Matrices by Hoteling's T-Square*

The existing similarity between different isolated dominant failure modes is calculated using Hotelling's T-square test. The data sets used for parity of failure modes are high dimensional. Data compression and de-correlation is performed for parity in 3-dimensional feature space. Hoteling's T-square test is used for pairwise comparison of various shock features belonging to different classes (failure modes) in a multivariate framework. Test statistic for this test is given by,

$$T^2 = \sqrt{n}(\bar{x} - \mu_0)(S)^{-1}(\bar{x} - \mu_0) \approx \frac{(n-1)p}{n-p} F_{\alpha}(p, n-p) \quad (29)$$

Where, 'n' is the no of observation in each class and 'p' signifies dimensionality of the observation. The above T-square formula, tests, if the mean of the class $\bar{x}$ is equal to μ_0 or not. The T-square statistic follows a scaled f-distribution and takes into account the covariance 'S' of the data set. Table 5(a) and Table 5(b) represent the similarity between different failure modes in the PoP test assembly for simulation and experimental data sets respectively. Values less than 0.05 indicate statistical dissimilarity between the two classes. Results indicate that different failure modes are dissimilar and that the main diagonal terms which are intended to be similar show a p-value of 1.

Table 5: Similarity Matrix for different fault modes in Error Seeded Data set.

	P	IF	IM	DC	DD	FC
P	1	0.23	0.18	0.21	0.37	0.12
IF		1	0.35	0.42	0.11	0.25
IM			1	0.15	0.20	0.14
DC		Diagonal Matrix		1	0.19	0.38
DD		Significance Level 90%			1	0.25
FC						1

Table 6: Similarity Matrix for different fault modes in Experimental Data set.

	P	IF	IM	DC	DD	FC
P	1	0.20	0.11	0.39	0.40	0.14
IF		1	0.28	0.34	0.27	0.68
IM			1	0.44	0.53	0.57
DC		Diagonal Matrix		1	0.60	0.36
DD		Significance Level 90%			1	0.25
FC						1

Table 7: Output combinations of neurons for inference on fault initiation.

Neuron Number								Neuron Output
6	1	0	0	0	0	0	0	
7	0	1	0	0	0	0	0	
8	0	0	1	0	0	0	0	
9	0	0	0	1	0	0	0	
10	0	0	0	0	1	0	0	
11	0	0	0	0	0	1	0	
Anamoly	P	IM	IF	DC	DD	FC	Mix-Mode	

C. *Inference on Damage Initiation and Progression*

In this work, prognostics and health monitoring decision framework has been developed for Package-on-Package component under accidental drop and shock. The pattern classifier using supervised neural net developed in this work can also be used for addressing damage diagnostics along with fault mode classification. The neural net can be used for tracking type of failure mode and localization of anomaly in the component.

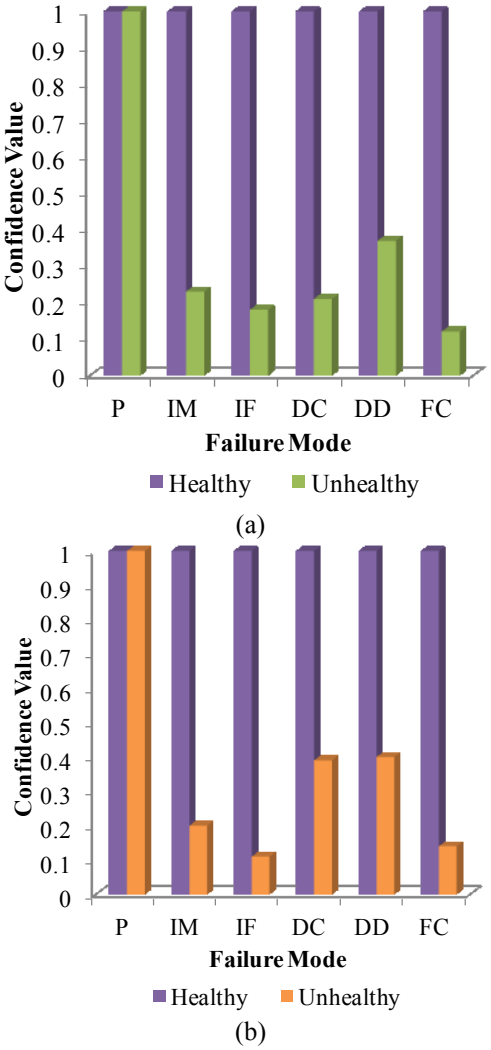


Figure 26: Inference on damage initiation and progression from statistical hypothesis. (a) Error Seeded Feature Space (b) Experimental Feature Space.

Table 7 summarizes the various possible outputs if the neural network. If this network is simulated with first three principal components of different shock feature then the output of various neurons in hidden layer-2 will fire output of various combinations as given in Table 7. For example if fifth column of Table 7 signifies that the PoP assembly has die cracking as the dominant failure mode, the neuron representing a die cracking state fires an output of 1 while other neurons remain dormant with an output of 0. This combination of output from fifth column is indicative that die cracking mode has initiated and is propagating. The localization of the damage is in vicinity of die region inside the PoP package. Hence this way fault initiation and

localization inside the component can be performed. Similarly as PoP assembly is subjected to subsequent shock events in time, the shock features in 3 dimensions based on time-frequency analysis can be used to simulate the network and inference on the state of damage and its location can be made.

Inference purely on damage initiation can also be performed by studying similarity matrices as given by Table 5. Figure 26a and Figure 26b represents the charts for the pair-wise comparisons of different shock features with the pristine state. A constant degradation in the confidence values is representative of accrued damage in the PoP test assembly when subjected to shock events. Hence damage initiation and progression thresholds are characterized using dissimilarities obtained from comparison of various shock features with pristine state.

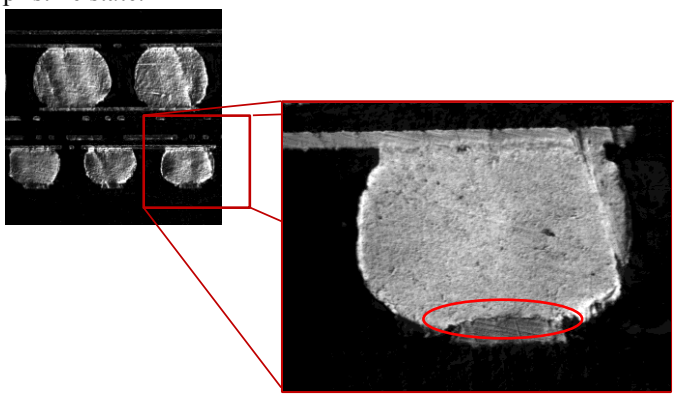


Figure 27: Experimental Cross sections of 2nd level Inter connects with cracking and fracture.

D. *Experimental Cross Section*

The failure modes classified in this study on PoP test assembly are validated using experimental cross sections of the failed PoP components. From the cross sections it is seen that significant cracking of the 2nd level interconnects takes place in the lower tier of the PoP package. Previous experience of the authors [Lall 2008, 2009, 20010, 2011] also suggests that cracking of 2nd level interconnects is one of the most dominant failure modes in electronic assemblies under drop and shock. Solder cracking as the most dominant failure mode, as gathered from visual inspection it can be seen that region representing solder cracking and failure is maximum in the volume. Figure 27 represents cracking of lower tier corner solder interconnects. The failure is seen on board and package sides.

X. SUMMARY

In this paper, a prognostics and health monitoring decision framework is presented for characterization of damage initiation and progression along with isolation of dominant failure modes for a Package-on-Package component-test assembly is presented. The PoP assembly has been studied under drop and shock environments. Transient dynamic strain histories obtained from the PoP assembly using strain-sensors and digital image correlation in conjunction with high-speed imaging. Error seeded simulation data set has been formed using explicit finite element simulations. The time-domain strain histories from experiment and simulation have been used for deriving feature vectors from time-frequency analysis. Karhunen-Loéve transform is used for de-correlation of the feature space. A 3-dimensional feature space has been used for hard parity of potential failure modes. Multilayer perceptron network has been designed for tracking of dominant failure modes. The classified failure modes are validated using multivariate analysis of variance and Hoteling's T-square. The failure analysis of PoP component is also performed under SEM for location of failure sites. The methodology presented in this work is scalable to system level reliability and can address damage diagnostics and fault mode classification in 3D packaging architectures such as Package on Package.

XI. ACKNOWLEDGMENTS

The research presented in this paper has been supported by members of the CAVE3 Electronics Research Center

REFERENCES

[1] Bishop, C. M., Neural networks for pattern recognition. Oxford University Press Inc., New York, USA, 1995.

[2] Cohen, L., Time-Frequency Analysis, Prentice-Hall, Englewood Cliffs,. NJ, 1995.

[3] Cohen, L., Time-Frequency Distributions-a Review, Proceedings of the IEEE, Vol. 77, Issue 7, pp. 941-981, 1989.

[4] Cheah, B.E., Kong, J., Periaman, S., Ooi, K. C., A Novel Inter-Package Connection for Advanced Package-on-Package Enabling, 61st Electronic Components and Technology Conference (ECTC) , Lake Buena Vista, Florida, pp. 589-594, May 31 - Jun 03, 2011.

[5] Das, R. N., Egitto, F. D., Bonitz, B., Poliks, M.D., Markovich, V.R., Package-Interposer-Package (PIP): A Breakthrough Package-on-Package (PoP) Technology for High End Electronics, 61st Electronic Components and Technology Conference (ECTC) , Lake Buena Vista, Florida, pp. 619-624, May 31 - Jun 03, 2011.

[6] Dreiza, M., L. Smith, G. Dunn, N. Vijayaragavan, J. Werner, Package on Package (PoP) Stacking and Board Level Reliability: Results of Joint Industry Study, Proceedings of IMAPS 2006.

[7] Georgopoulos, V.C., Preis, D., A Non-Negative Approximate Wigner Distribution With Accurate Low-Order Moments, International Conference on Digital Signal Processing Vol. 2, pp. 769 – 772, 1997.

[8] Grossmann, A., J. Morlet, Decomposition of Hardy Functions into Square-Integrable Wavelets of Constant Shape, SIAM J. Appl. Math., vol. 15, pp. 723-736, 1984.

[9] Haykin, S., Neural Networks: A Comprehensive Foundation, second edition, Prentice Hall, New Jersey, USA, 1999.

[10] Jain, A.K., Mao, J., Mohiuddin, M., Neural Networks: A Tutorial, IEEE Computer, Vol.29, 3, pp. 31-44, March 1996.

[11] Johnson, R. A., Wichern, D. W., Applied Multivariate Statistical Analysis, 6th Edition, Prentice Hall, 2007.

[12] Jeong, J., Williams, W.J., Alias-free Generalized Discrete-Time Time-Frequency Distributions, IEEE Transactions on Signal Processing, Vol. 40, Issue 11, pp. 2757 – 2765, 1992a.

[13] Jeong, J., Williams, W.J., Kernel Design For Reduced Interference Distributions, IEEE Transactions on Signal Processing, Vol. 40, Issue 2, pp. 402 – 412, 1992b.

[14] Lall, P., Panchagade, D., Liu, Y., Johnson, W., and Suhling, J., Models for Reliability Prediction of Fine-Pitch BGAs and CSPs in Shock and Drop-Impact, 54th Electronic Components and Technology Conference, pp. 1296–1303, 2004.

[15] Lall, P., Choudhary, P., Gupte, S., Health Monitoring for Damage Initiation and Progression during Mechanical Shock in Electronic Assemblies, 56th Electronic Components and Technology Conference, San Diego, California, pp.85-94, May 30-June 2, 2006a.

[16] Lall, P., Panchagade, D., Liu, Y., Johnson, W., Suhling, J., Models for Reliability Prediction of Fine-Pitch BGAs and CSPs in Shock and Drop-Impact, IEEE Transactions on Components and Packaging Technologies, Volume 29, No. 3, pp. 464-474, September 2006b.

[17] Lall, P., Choudhary, P., Gupte, S., Suhling, J., Hofmeister , J., Statistical Pattern Recognition and Built-in Reliability Test for Feature Extraction and Health Monitoring of Electronics under Shock Loads, 57th Electronic Components and Technology Conference, Reno, Nevada, USA, pp. 1161-1178, May 29-June 1, 2007a.

[18] Lall, P., Panchagade, D., Iyengar, D., Shantaram, S., Suhling, J., Schrier, H., High Speed Digital Image Correlation for Transient-Shock Reliability of Electronics, 57th Electronic Components and Technology Conference, Reno, Nevada, pp. 924-939, May 29 – June 1, 2007b.

[19] Lall, P., Gupta, P., Kulkarni, M., Panchagade, D., Suhling, J., Hofmeister , J., Time-Frequency and Auto-Regressive Techniques for Prognostication of Shock-Impact Reliability of Implantable Biological Electronic Systems, 58th Electronic Components and Technology Conference, Orlando, Florida, pp. 1196-1207, May 27-30, 2008.

[20] Lall, P., Gupta, P., Panchagade, D., Angral, A., Fault-Mode Classification for Health Monitoring of Electronics Subjected to Drop and Shock, 59th Electronic Components and Technology Conference, San Diego, California USA, pp. 668-681, May 25-29, 2009.

[21] Lall, P., Gupta, P., Panchagade, D., Self-Organized Mapping of Failure Modes in Portable Electronics Subjected to Drop and Shock, Proceeding of 60th Electronic Components & Technology Conference, Las Vegas, NV, pp. 1195-1208, June 1-4, 2010.

[22] Lall, P., Gupta, P., Goebel, K., Identification of Failure Modes in Portable Electronics Subjected to Mechanical-Shock using Supervised Learning of Damage Progression, 61st Electronic Components and Technology Conference (ECTC) , Lake Buena Vista, Florida, pp.1944-1957, May 31 - Jun 03, 2011.

[23] Lall, P., Gupta, P., Goebel, K., Decorrelated Feature Space and Neural Nets Based Framework For Failure Modes Clustering in Electronics Subjected to Mechanical-Shock, IEEE International Conference on Prognostics and Health Management (PHM), Denver, Colorado, June 20-23, 2011.

[24] Lall, P., Angral, A., Suhling, J., Reliability Studies for Package-on-Package Component in Drop and Shock Environments, Technical Conference & Exposition on Packaging and Integration of Electronic and Photonic Systems, MEMS and NEMS INTERPACK'11, Portland, Oregon, pp. 1-12, July 6-8, 2011.

[25] Lin, W., Yoshida, A., Dreiza, M., Control of the Warpage For Package-on-Package (PoP) Design, SMTA International Conference, pp.320-326, 2006.

[26] Lin, J.C., Yu, J., Chung, B., Chang, K., Fang, D., Development on Ultra High Density Memory Package with PoP Structure, 61st Electronic Components and Technology Conference (ECTC) , Lake Buena Vista, Florida, pp. 1136-1140, May 31 - Jun 03, 2011.

[27] Ogawa, H., Oja, E., Projection filter, Wiener filter and Karhunen Loeve subspaces in digital image restoration, Journal of Math. Anal. Appl. Vol 114 , no. 1, pp. 37-51, 1986.

[28] Rumelhart, D. E., Hinton, G. E., Williams, R. J., Learning Representations by Back -Propagating Errors, Nature, vol. 323, pp. 533-536, 1986.

[29] Schalkoff, R., Pattern Recognition: Statistical, Structural, and Neural Approaches. New York, Wiley, 1992.

[30] Shawn, K., Anderson, R.H., Krolik, J. L., A Generalized Karhunen-Loeve Basis for Efficient Estimation of Tropospheric Refractivity Using Radar Clutter, IEEE Transactions on Signal Processing, vol. 52, no. 1, pp. 48-60, 2004.

[31] Tacer, B., Loughlin, P.J., Instantaneous Frequency and Time-Frequency Distributions, International Conference on Acoustics, Speech, and Signal Processing, Vol. 2, pp. 1013-1016, 1995.

[32] Tacer, B., Loughlin, P.J., What are the Joint Time-Frequency Moments of a Signal?, International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, pp. 1427-1430, 1996.

[33] Ville, J., Theorie et Applications de la Notion de Signal Analytique, Cablet et Transmission, Vol. 2A, pp. 61-74, 1948.

[34] Wigner, E., On the Quantum Correction for Thermodynamic Equilibrium, Physics Rev, Vol. 40, pp. 749, 1932.

[35] Williams, W.J., Tzuhsien, S., Adaptive RID Kernels Which Minimize Time-Frequency Uncertainty, IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis, pp. 96-99, 1994.

[36] Wilamowski., B. M., Neural Networks and Fuzzy Systems for Nonlinear Applications, 11th International Conference on Intelligent Engineering Systems, 29 June - 1 July 2007, Budapest, Hungary.

[37] Yamashita, Y., Ogawa, H., Relative Karhunen Loeve Transform, IEEE Transactions on Signal Processing, vol. 44, no.2, pp. 371-378, 1996.

[38] Yim, M. J., Strode, R., Brand, J., Adimula, R., Zhang, J. J., Yoo, C., Ultra Thin POP Top Package using Compression Mold: Its Warpage Control, 61st Electronic Components and Technology Conference (ECTC) , Lake Buena Vista, Florida, pp. 1141-1146, May 31 - Jun 03, 2011.

[39] Yoshida, A., Taniguchi, J., Murata, K., Kada, M., Yamamoto, Y., Takagi, Y., Notomi, T., Fujita, A., A Study on Package Stacking Process for Package-on-Package (PoP), 56th Electronic Components and Technology Conference (ECTC) , San Diego, California, pp. 6, May 30 - Jun 02, 2006.