Prognostic Health Management for Mission and Safety Critical Electronics

by

Ryan D. Lowe

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Auburn, Alabama
August 4, 2012

Keywords: Prognostic Health Management, Electronics, Prognostics,
Health Monitoring, Kalman Filter, Lead Free Solder

Approved by

Pradeep Lall, Chair, T. Walter Professor of Mechanical Engineering
Jeffrey C. Suhling, Quina Distinguished Professor of Mechanical Engineering
George T. Flowers, Professor of Mechanical Engineering
Michael J. Bozack, Professor of Physics

Abstract

The increasing reliance on electrical systems to fulfill mission and safety critical applications has motivated the need for *in situ* monitoring of functioning electrical systems and *a priori* warnings of failure. In this research the drawbacks of traditional reliability methods when applied to mission and safety critical applications are discussed and a new paradigm for the reliability of electronics, prognostic health management, is demonstrated. In the prognostic health management reliability model wear and damage to individual electrical systems are monitored and advanced warnings of failure are issued to allow adequate time to plan mitigating action and avoid unplanned failures. The occurrence of an unplanned failure in a mission and safety critical electrical component is considered by definition to have an associated cost that is unbearably high. Leading indicators of failure have been developed that allow in-situ monitoring of the structural health of electronics, a method coined *resistance spectroscopy*. The presented techniques are non-destructive in nature and were purposefully designed to cheaply embed into new electrical systems. Methods for processing the stream of information from real time observations are presented in a manner that facilitates statistically defendable decision making and the optimization of safety, availability, and operating costs. A variety of recursive filters – least squares, Kalman, extended Kalman, and particle – are combined with prognostic methods (forecasting) to create seamless real time monitoring and prediction algorithms. Solder joint configurations studied include SnPb eutectic, high lead, lead-free (SnAgCu), copper columns, and micro coil springs. Architectures studied include ball grid array, land grid array, a novel

spring interconnect and pin/spring electrical connectors. Test environments include drop/shock, vibration, and simultaneous temperature and vibration.

Investigations relating to the implementation of the presented techniques demonstrate the practical nature of the work. A cost justification method that is accessible to engineers, technical managers, and executives is developed to quantify the business case for implementing prognostic health management for electronics. Particle swarm optimization methods have been used to demonstrate the expected future performance of implemented prognostic algorithms given a set of test data. The described methods have been shown to be particularly sensitive to damage in the novel spring interconnect designed for long duration space travel.

Methods for verifying the correct operation of prognostic algorithms and validating that algorithms meet specified requirements (both online and offline) are discussed. A prognostic health management toolbox, coded in Matlab™, has been created. The toolbox provides a foundation of verified and validated algorithms for implementation of the presented methods, and is highly extendable for the development of new prognostic algorithms. The code is generic enough to apply to electronics as well as other application domains where suitable leading indicators of failure exist. All of the work presented in this document was created in part with the prognostic health management toolbox.

Acknowledgements

I would like to acknowledge the love and support of my wife, who without the slightest qualm worked multiple jobs to support me during my time in graduate school. I cannot think of a single time when she complained about the strains that my education placed on her. My parents and siblings have helped me endure the ups and downs that accompany graduate school and facilitated my escape for vacations in the mountains. My in-laws have been a source of love, support, and nutrition in Auburn. I am very fortunate to have such a supportive family. I would also like to thank the Madsen's for treating me as if I was family during our time in Auburn.

Academically Auburn has been a prosperous place for me. I hope to take from Auburn the insightful teaching methods that Dr. Suhling showcased in his many classes. If only everything could be made to seem so easy. To my wife's dismay, I found the joy of building mechanical toys, such as double pendulums, as a result of attending Dr. Flowers class. Dr. Bozack has helped reinforce the attitude that ANYTHING can be fixed, his mechanical aptitude is beyond description. I would have never known what a vacuum tube was if it was not for Dr. Bozack! For fear of missing someone, I will not list every one of my lab mates who provided countless hours of training, support, humor, and technical feedback. I am honored to say that I crossed paths with many of you, thank you for sharing your lives with me.

I would also like to acknowledge the environment that my advisor, Dr. Lall, fostered at the Center for Advanced Vehicle and Extreme Environment Electronics ($CAVE^3$). I can honestly say that I had fun working in his lab. His unique mixture of technical skill, craftsmanship,

persona, and a genuine intellectual curiosity is a rare combination. I already miss the fast paced and exciting research that he is immersed in.

Table of Contents

Abstract……..……..……ii
Acknowledgements……..iv
List of Figures……..ix
List of Tables……..…..…xviii
1 Introduction 1
1.1 Mission and Safety Critical Electronics 1
1.2 Traditional Reliability Approaches for Electronics 2
1.3 Shortcoming of Traditional Life Prediction Approaches 4
1.4 Prognostic Health Management 6
1.5 Risk Based Decision Making 9
1.6 Resistance Spectroscopy 9
1.7 Overview of Electronic Packaging 10
2 Literature Review 15
2.1 Traditional Reliability Methods 15
2.2 Resistance Spectroscopy 23
2.3 Prognostic Health Management 23
3 Resistance Spectroscopy 33
3.1 Direct Current Wheatstone Bridge 33
3.2 Alternating Current Wheatstone Bridge 39
3.3 Phase Sensitive Detection 43
3.4 Theory vs. Observation at Different Frequencies 46
4 Risk Based Decision Making 49
4.1 Methodology 49
4.2 Conclusion 54
5 Demonstration of Recursive Filtering 56
5.1 Least Squares Approaches 56
5.2 State Estimator Approaches 68
5.3 Tracking Example: Physics of a Ballistic Object Re-entry Path 88
5.4 Conclusion 111
6 Quantifying performance of PHM implementations 112
6.1 Example PHM Implementation 112
6.2 Prognostics Hindecaste 116
6.3 PHM Specific Offline Performance Metrics 123
6.4 Conclusion 130
7 Prognostics with Resistance Spectroscopy for Grid Array Packaging 131
7.1 Test Vehicle 132
7.2 Resistance Spectroscopy Approach 133
7.3 The Effect of Shear Strain on Change in Resistance 134
7.4 Phase Sensitive Detection 138
7.5 Experimental Setup 140

7.6 Training Signal and Identification of Onset of Failure 141
7.7 Prognostic Distance 144
7.8 Failure Modes 149
7.9 Conclusion 151
8 Recursive Least Squares (RLS) Filtering Based Prognostic Algorithm Implemented for Electrical Connectors 152
8.1 Introduction 152
8.2 Test Setup 153
8.3 Resistance Monitoring 155
8.4 Prognostic Hindecaste 157
8.5 Conclusion 166
9 Kalman Filter (KF) Based Prognostic Algorithm Implemented for BGA's 167
9.1 Introduction 167
9.2 Test Vehicle 168
9.3 Transfer Function for Interconnect Strain to Resistance 171
9.4 Filtering and RUL Prediction 177
9.5 Estimation of Remaining Useful Life 182
9.6 Sensitivity Study 187
9.7 Risk Based Decision Making 189
9.8 Conclusion 191
10 Extended Kalman Filter (EKF) Based Prognostic Algorithm Implemented for BGA's 192
10.1 Introduction 192
10.2 Test Vehicle 193
10.3 Filtering and RUL prediction 196
10.4 Prognostic Hindecaste 200
10.5 PHM Metrics 205
10.6 Comparison Against the Kalman Filter 206
11 Particle Filter (PF) Based Prognostic Algorithm Implemented for BGA's 210
11.1 Introduction 210
11.2 Test Vehicle 211
11.3 Test Environment 212
11.4 Test Circuitry 213
11.5 Filtering and RUL Predictions 218
11.6 Prognostic Metrics 225
11.7 Comparison Against the Extended Kalman Filter 227
11.8 Conclusion 229
12 PHM for electronics subjected to multiple simultaneous stress environments 231
12.1 Introduction 231
12.2 Test Vehicle 234
12.3 Test Environment 235
12.4 Test Circuitry 237
12.5 Simultaneous Temperature and Vibration Testing 240
12.6 Prognostic Capability 244
12.7 Critical Resistance 245
12.8 Conclusion 246
13 Particle Swarm Optimization for PHM 248
13.1 Introduction 248
13.2 Particle Swarm Optimization 250
13.3 Conclusion 257

14 Cost justification for Implementing PHM on circuit boards 259
14.1 Introduction 260
14.2 The Real Option for R&D Investment in Prognostic Technology 262
14.3 Black-Scholes Formula 263
14.4 Datar-Mathews Approach 265
14.5 Equivalence of Black-Scholes and Datar-Mathews Approach 266
14.6 Calculating the Value of the Option to Invest in PHM 270
14.7 Comparison of Real Option Analysis with Net Present Value Approach 278
14.8 Comparison with Decision Tree Modeling Approach 279
14.9 Convergence of Monte Carlo Simulations 280
14.10 Conclusion 281
14.11 Appendix 281
15 PHM for the Micro-Coil Spring Interconnect 287
15.1 Introduction 287
15.2 Test Vehicle 289
15.3 Experimental Setup 291
15.4 Health Monitoring Circuit 293
15.5 Health Monitoring Algorithm 298
15.6 Prognostics Hindecaste 300
15.7 Verification and Validation 310
15.8 Filter Divergence 315
15.9 Skill Against Weibull Baseline 316
15.10 Conclusion 318
15.11 Appendix 319
16 PHM Matlab Toolbox 325
16.1 Introduction 325
16.2 Verification and Validation Scripts 326
16.3 Code documentation 327
17 Conclusion 329
17.1 Salient Features of the Presented Work 329
17.2 Future Work 333
18 References 336

List of Figures

Figure 1: Hypothetical failure distribution highlighting some drawbacks of traditional reliability methods 8
Figure 2: Ground truth PDF used for the Weibull accuracy simulation 18
Figure 3: Error in MLE parameter estimation for different sample sizes compared to the ground truth simulation 18
Figure 4: Error in MLE parameter estimation for different sample sizes compared to the ground truth simulation 19
Figure 5: DC Wheatstone bridge 33
Figure 6: Equivalent circuit for a DC Wheatstone bridge 38
Figure 7: Addition of a shunt resistor to facilitate measurement of changes in resistance of the small resistance Rx 38
Figure 8: Sensitivity of the DC Wheatstone bridge for different choices of resistors 39
Figure 9: Schematic of AC Wheatstone Bridge 40
Figure 10: Bode plot for AC Wheatstone bridge 43
Figure 11: Phasor representation of a sine wave 45
Figure 12: Functional Schematic of Lock in Amplifier 45
Figure 13: Simplified block diagram adapted from SRS830 Lock-in amplifier user manual 46
Figure 14: Bode plot from a circuit optimized for a low frequency measurement 47
Figure 15: Magnitude component recorded at different driving frequencies for the same test 48
Figure 16: Zoomed view aruond the region of interest from Figure 15 48
Figure 17: Typical wear out profile of a safety and mission critical system 49
Figure 18: Failure threshold and detection threshold that facilitates quantifying failures based on the value of a feature vector 50
Figure 19: The prediction challenge, estimating the time until the feature vector crosses the failure threshold 51
Figure 20: Uncertainty in the failure prediction 51
Figure 21: Window of opportunity for taking action 53
Figure 22: Appropriate time to schedule maintenance 54
Figure 23: Updated failure prediction after new information is available 54
Figure 24: Simulation of a constant signal at a value of 1.2 measured in the presence of white noise 59
Figure 25: Comparison of the batch least squares approach to the recursive least squares approach 60
Figure 26: Simulated noisy data with an intercept of 1.2 and a slope of 2 61
Figure 27: Batch least squares fit of noisy data 62
Figure 28: Comparison of batch least squares to recursive least squares for the a0 polynomial term 62
Figure 29: Comparison of batch least squares to recursive least squares for the a1 polynomial term 63

Figure 30: Synthesized spectrometric data set denoted as (Clean) and a noise corrupted version of the same signal (Noisy) 65
Figure 31: Moving average filter of length 65 applied to the synthetic data set 65
Figure 32: Savitzky-Golay smoothing applied to the noisy data set with a window length of 65 and a model order of 10 66
Figure 33: Numeric derivatives estimated for a sinusoidal signal corrupted with white noise 67
Figure 34: Savitzky-Golay estimates of derivatives of noisy data with a window of 41 and a model order of 3 67
Figure 35: Approximation of an ideal and continuous probability density function by a set of particles. Particles locations represented as xki and specify the position on the x-axis. Weights specify the height of the stems and are represented by wki 84
Figure 36: a) initial probability mass function represented by five particles, b) cumulative distribution function, c) assigning draws from uniform distribution U(0,1) to new particle locations based on cumulative distribution function, d) importance re-sampled distribution 87
Figure 37: Evolution of Model Prediction in Particle Filter 88
Figure 38: Simple example problem for generating a baseline data set 89
Figure 39: Simulated position of ballistic object 90
Figure 40: Simulated velocity of ballistic object 90
Figure 41: Simulated acceleration of ballistic object 91
Figure 42: Effect of drag coefficient on simulation 91
Figure 43: Simulating a noisy measurement from the ground truth simulation. The simulation line is essentially the same, and therefore not visible in the plot, as the noisy measurement at this level of resolution 92
Figure 44: Zoomed view of the noisy measurement 92
Figure 45: Tracking results for a two term PVA Kalman filter 94
Figure 46: Residual of the error between the estimated and simulated (no noise) states 95
Figure 47: Benchmark Testing for Various Implementations of recursive filters 100
Figure 48: Tracking results for EKF PVB filter with a poorly initialized ballistic coefficient .. 103
Figure 49: Tracking results for PF PVB filter with a poorly initialized ballistic coefficient 104
Figure 50: Tracking performance for PVB filters with the beta parameter initialized correctly and incorrectly 104
Figure 51: Estimate of ballistic coefficient. Note the very fine scale on the y-axis. 105
Figure 52: Simulation with the added complexity of a time varying drag coefficient 107
Figure 53: Change in the simulated beta coefficient. The value is reduced by 50% for the final 15 seconds of the simulation. 107
Figure 54: Tracking performance for the KF with an evolving beta parameter 107
Figure 55: Tracking performance for the EKF with an evolving beta parameter 108
Figure 56: Tracking performance for the PF with an evolving beta parameter 108
Figure 57: Tracking performance for a poorly initialized EKF with an evolving beta parameter 109
Figure 58: Tracking performance for a poorly initialized PF with an evolving beta parameter 109
Figure 59: Comparison of recursive filters for a time evolving beta parameter 110
Figure 60: Filter estimate of the drag coefficient 111
Figure 61: Noisy measurement. Units are purposefully omitted to avoid accidently mistaking this data as a real experimental observation. 113

Figure 62: Tracking results from the Kalman filter for the synthetic data set 115
Figure 63: The Kalman filter successfully tracked the noisy data and provides smooth estimates of the first two derivatives. As initially assumed, the second derivative is approximately constant 115
Figure 64: Residuals between the estimated and actual state estimates falls within the theoretical error bounds 66% of the time and indicates that the Kalman filter is operating correctly. Unfortunately this provides no insight if RUL predictions are accurate and correct 116
Figure 65: Illustration of the Prediction Process with an Overlay on the reference data set. With little trend in the measured data provides little insight into the nature in which damage will propagate in the system 117
Figure 66: Prediction after 50 measurements. With an observation of available data that is not above the noise floor, the prediction of RUL is significantly over estimated 118
Figure 67: Prediction after 75 measurements. At this point in time enough measurements have been made to more accurately estimate the form of the underlying process that is generating the noisy data. The algorithm is still slightly overestimating the actual RUL 118
Figure 68: Prediction after 100 measurements. In contrast to the previous predictions, at this point in time the RUL is slightly under predicted. 119
Figure 69: Prediction after 125 measurements. This prediction represents the first prediction that has a very high accuracy compared to the actual RUL 119
Figure 70: Prediction after 150 measurements. At this point in the simulation it is apparent that the PHM algorithm has converged to the actual RUL. 120
Figure 71: Prediction after 175 measurements. Even though the PHM algorithm has converged to a accurate prediction updates are still made in the event that the underlying process is affected by an un-modeled input 120
Figure 72: Prediction after 200 measurements. During the final stages of the simulation the fictitious product is incurring damage at a much faster rate. 121
Figure 73: Prediction after 225 measurements. The final prediction before failure occurs. For such short time scales the specific shape of the failure propagation curve is no longer important. A linear model would perform nearly as well as a quadratic model. 121
Figure 74: An illustrative example to illustrate how RUL plots are calculated 122
Figure 75: Summary of RUL predictions for the hindcaste on the fictitious data set 123
Figure 76: Alpha-Lambda plot which shows both RUL predictions and the uncertainty for each prediction 124
Figure 77: Illustration of the Beta calculation for a Gaussian distribution at a prediction 60% of the way to failure 126
Figure 78: Beta metric for the simulated data set. Note the decrease in the metric score near the actual time of failure 127
Figure 79: Numeric integration could be used to evaluate the area between two alpha bounds defined by the probability mass function denoted as the approximation to the blue ideal curve. This situation is common in a particle filter implementation base PHM algorithm. 128
Figure 80: Relative accuracy for the simulated PHM predictions. Note the sharp drop off in relative accuracy towards the end of life. 129
Figure 81: Test Vehicle 132
Figure 82: Schematic representation of solder ball denoting length and cross sectional area 134

Figure 83: Hypothesized effect of PCB bending on solder joint interconnects 135
Figure 84: Mesh convergence study for the solder joint geometry under an axial loading 136
Figure 85: Axial strain case with the top of the solder joint constrained and the bottom of the solder joint is displaced downward. The grid represents the unreformed shape, and the solid represents the deformed shape 136
Figure 86: Shear strain case with the top of the solder joint constrained and the bottom of the solder joint is displaced to the right. The grid represents the unreformed shape, and the solid represents the deformed shape 137
Figure 87: Comparison of trends in the calculated change of resistance for the two simulations 137
Figure 88: The simulated state of shear, shown on the left, is not a pure state of shear and adds small amounts of axial strain to the simulation. A pure state of shear creates no change in length of a strain element and is shown on the right. The pure state of shear is difficult to apply to a solder joint geometry 138
Figure 89: Wheatstone Bridge with capacitors C1, C2 and resistors R1, R2, R3, PKG 139
Figure 90: Lansmont Model 23 Shock Test System 140
Figure 91: Failure Metric Calculation from Electrical Continuity. 141
Figure 92: Phase shift of healthy package (Test Board-A, SAC305) 141
Figure 93: Repeatability of phase shift measurement on pristine healthy packages (Test Board-A, SAC alloy) 142
Figure 94: Confidence value as a lead indicator of failure during a drop test (Test Board-A, SAC alloy) 143
Figure 95: Degradation of confidence value during drop test (Test Board-A, CBGA Package U2, f = 6MHz) 146
Figure 96: Degradation of confidence value during drop test (Test Board-A, f = 6MHz) 147
Figure 97: Degradation of confidence value during drop test of packages U4,U6 and U7 on for all interconnects. U4 is shown with blue circles, U5 is shown with green squares and U6 is shown with red crosses (Test Board-A, 127 kHz) 148
Figure 98: Degradation of confidence value during drop test of packages U4,U6 and U7 on a normalized scale for all interconnects. U4 is shown with blue circles, U5 is shown with green squares and U6 is shown with red crosses (Test Board-A, 6MHz) 148
Figure 99: Method for determining prognostic distance using a threshold value shown in red. Each trace is an individual package. (Test Board-A, SAC305 alloy) 149
Figure 100: PDF of prognostic distance at varying frequencies. (Test Board-A, SAC alloy) ... 149
Figure 101: Damaged SAC interconnect (Test Board A) 150
Figure 102: Broken trace on EUT interconnect (Test Board A) 150
Figure 103: Cracks through HIPB interconnect (Test Board A) 150
Figure 104: Failed CCGA interconnect (Test Board A) 151
Figure 105: Diagram of Experimental Setup 153
Figure 106: Diagram showing the internal configuration of the connector. Relative motion between the pin (rigidly attached to the header) and the spring (rigidly attached to the housing) results in fretting at the mating interface. 154
Figure 107: Fixture used to apply input to the connector header 154
Figure 108: Zoomed view of the fixture used to apply vibration inputs to the connector header 155

Figure 109: During testing the shaker table vibrates in the vertical direction (right side of picture). The motion of the shaker table has caused the image of the connector to blur. The rigid mechanical connection on the left shows no motion and did not blur in the photograph 155
Figure 110: Time history of two wire resistance measurements and the increase in contact resistance caused by connector fretting 156
Figure 111: Contact surface of the spring before testing 156
Figure 112: Deep grooves and asperities on a tested spring caused by relative motion between the spring and pin 157
Figure 113: Batch least squares fit of a linear model to the experimentally observed data 158
Figure 114: Recursive least squares state estimates 160
Figure 115: Remaining useful life predictions 161
Figure 116: Comparison of batch least squares to recursive least squares estimates for the slope of the best fit line. Note that the batch and recursive estimate for the slope of the best fit line are identical for the full data set. 161
Figure 117: Prediction snapshot from early in the test. The lack of trend in the data results in an RUL prediction that is overly optimistic 162
Figure 118: Further into the test RUL predictions are improving but still have not converged. 162
Figure 119: At this point in the test sufficient data is available to estimate the slope of a best fit line through the data 163
Figure 120: At the end of the test the estimate of the slope of the best fit line is nearly identical the RLS fit 163
Figure 121: Alpha-lambda plot summarizing the uncertainty in the RUL predictions 164
Figure 122: Time history of the uncertainty reported by the PHM algorithm. Note the log scale on the y-axis. 164
Figure 123: Beta metric calculation for the PHM hindecaste 165
Figure 124: Relative accuracy metric for the PHM hindecaste 165
Figure 125: Test Board 169
Figure 126: Step stress profile for vibration testing that fatigues interconnects to failure. 170
Figure 127: Random vibration profile at varying g levels corresponding to the step stress profile outlined in Figure 126 170
Figure 128: Raw resistance data. The data used as an input data vector is shown in the brackets 171
Figure 129: Zoomed view of resistance data between 2 hrs and failure 172
Figure 130: Constraints on solder ball for FEM simulation 174
Figure 131: Meshed model of solder ball 175
Figure 132: Deformed and undeformed geometry of solder ball 175
Figure 133: Simulated change in resistance of solder ball during pull test. Arrows indicate expected change in resistance at a strain of 0.1. 176
Figure 134:Graphical state space representation of a system 178
Figure 135: Results of Kalman Filtering 183
Figure 136: Convergence of the Kalman gain for the resistance estimate term. 183
Figure 137: RUL prediction at 2.6 Hrs, the red circle shows what data was available for the prediction. The blue line in both plots is the feature vector, and the green line is the extrapolated state value used to predict RUL 184
Figure 138: RUL prediction at 3.8 Hrs 184

Figure 139: RUL prediction at 5.6 Hrs 184
Figure 140: Comparison of actual RUL vs predicted RUL 185
Figure 141: Alpha-Lambda accuracy metric illustrating the uncertainty surrounding RUL predictions 186
Figure 142: Beta metric illustrating the percentage of the RUL PDF that overlapped with the alpha bounds for each prediction 186
Figure 143: Relative Accuracy of the RUL predictions 187
Figure 144: Variation in the sum of the beta calculation for variations in the critical-threshold of state variable. 188
Figure 145: Variation in the sum of beta calculation for variations in tunable the process noise parameter 188
Figure 146: Time to order replacement component calculation vs time, which demonstrates one method of statistically defendable decision making using estimates of RUL. 190
Figure 147: Test Board 194
Figure 148: Step stress profile for vibration testing that fatigues interconnects to failure. 195
Figure 149: Zoomed view of resistance data between 2 hrs and failure 195
Figure 150: EKF filtering results 201
Figure 151: Convergence of Kalman gain 201
Figure 152: RUL prediction at 4.4069 Hrs, the red circle shows what data was available for the prediction. The blue line in both plots is the feature vector, and the green line is the extrapolated state value used to predict RUL 203
Figure 153: RUL prediction at 5.0564 Hrs 203
Figure 154: RUL prediction at 5.7059 Hrs 203
Figure 155: Comparison of actual RUL vs predicted RUL 204
Figure 156: Residual between the estimated one step ahead prediction and the noisy measurement 205
Figure 157: Alpha-Lambda Performance Metric 206
Figure 158: Beta Precision Metric 206
Figure 159: Relative Accuracy Metric 206
Figure 160: a) RUL plot for KF, b) RUL plot for EKF 208
Figure 161: a) Alpha-Lambda plot for KF, b) Alpha-Lambda plot for EKF 208
Figure 162: a) Beta metric for KF, b) beta metric for EKF 208
Figure 163: a) Relative accuracy for KF, b) Relative accuracy for EKF 209
Figure 164: PBGA324-1mm-19mm test board 211
Figure 165: Vibration profile used for testing 213
Figure 166: Functional diagram of test circuitry. The block labeled package represents the daisy chain resistance of the package under test 213
Figure 167: Resistance spectroscopy measurement setup. A differential output from the bridge is input into the lock-in amplifier for phase sensitive detection. Magnitude and phase data are recorded with the data logger. 215
Figure 168: Phase sensitive detection performed inside the lock-in amplifier uses digital signal processing to multiply two sinusoids together and then eliminate high frequency components with a filter 217
Figure 169: Change in resistance measured with the resistance spectroscopy method. Notice the significant change in resistance that is detectable well before the traditional definition of failure. 218

Figure 170: Theoretical relation between resistance, material properties, and geometry for a cylinder. A similar, more complicated relation exits for a solder interconnect 220
Figure 171: Proposed possible method of crack propagation in a solder interconnect resulting in a change in resistance. Electrical continuity is lost when crack length, a, reaches one. ... 221
Figure 172: Results of particle filtering applied to feature vector shown in Figure 169 223
Figure 173: Damage propagation to predict EOL with a particle population size of N = 50 224
Figure 174: Predicted end of life distributions reported at decreasing times to failure. Darker lines represent predictions closer to the true end of life. 225
Figure 175: Remaining useful life predictions 225
Figure 176: Alpha-Lambda Performance of PHM Algorithm 226
Figure 177: Beta calculation showing area under RUL prediction PDF that falls within the alpha bounds 226
Figure 178: Relative Accuracy of RUL prediction 227
Figure 179: a) RUL plot for PF, b) RUL plot for EKF 228
Figure 180: a) Alpha-Lambda plot for PF, b) Alpha-Lambda plot for EKF 228
Figure 181: a) Beta metric for PF, b) beta metric for EKF 229
Figure 182: a) Relative accuracy for PF, b) Relative accuracy for EKF 229
Figure 183: PBGA324 test vehicle 235
Figure 184: Vibration profile used for testing 236
Figure 185: Thermal chamber with upper window mounted on top of the shaker table 236
Figure 186: Deflection of the center of the board during vibration testing at different ambient temperatures 236
Figure 187: Functional diagram of test circuitry. The block labeled package represents the daisy chain resistance of the package under test 237
Figure 188: Resistance spectroscopy measurement setup. A differential output from the bridge is input into the lock-in amplifier for phase sensitive detection. Magnitude and phase data are recorded with the data logger. 239
Figure 189: Discrete component values used in resistance spectroscopy AC Wheatstone bridge 239
Figure 190: Phase sensitive detection performed inside the lock-in amplifier uses digital signal processing to multiply two sinusoids together and then eliminate high frequency components with a filter 240
Figure 191: Time history of daisy chain resistance for pristine boards tested at 25C 241
Figure 192: Time history of daisy chain resistance for pristine boards tested at 125C 242
Figure 193: Time history of daisy chain resistance for isothermally aged boards tested at 25C 242
Figure 194: Time history of daisy chain resistance for isothermally aged boards tested at 125C 243
Figure 195: Graphical representation of failure distributions from extracted Weibull parameters 243
Figure 196: Zoomed view of pristine-25C, pristine-125C, and aged-25C failure distributions. 244
Figure 197: Zoomed view of Figure 8 at the time just before failure occurs. The resistance spectroscopy technique is sensitive to a measurable change in the resistance of the packaged before the traditional IPC definition of failure 245
Figure 198: Additional tests run to create a test set for the PHM algorithm 249
Figure 199: Leave one out cross validation procedure repeated for each data set 249
Figure 200: Initialization of particles for two of the free parameters 252

Figure 201: Time history of swarm metrics. PSO terminates when the metrics converge below a specified tolerance. 255
Figure 202: Search space of the PSO algorithm visualized for the process noise and measurement noise terms in the Kalman filter 255
Figure 203: Comparison of PSO method against previously published results to show the improvement using the new method. 256
Figure 204: Performance of PHM algorithm when using only portion of the initial data set to find optimum parameters 257
Figure 205: Price index for 1948-2004 showing trends in agriculture inputs and outputs. Re-drawn from Fuglie, Keith O., James M. MacDonald, and Eldon Ball. Productivity Growth in U.S. Agriculture. EB-9, U.S. Dept. of Agriculture, Econ. Res. Serv. September 2007. 262
Figure 206: Hypothetical evolution in time of an investment using Equation (1). In this example the volatility is 50%. Discretely simulated values have been connected with lines to show trends 264
Figure 207: Comparison of a Monte Carlo simulation of the future price of an investment compared to the analytical result. The analytical solution is a lognormal distribution. The Monte Carlo simulation used 106 trials of Equation (1) binned at 60 locations. 265
Figure 208: Cash flow diagram for example problem 267
Figure 209: Convergence of the DM method to the Black-Scholes value for the example problem. Error bars represent one standard deviation determined from 30 different trials. 268
Figure 210: Three predicted levels of operating profits. These numbers are derived from a more detailed analysis shown in the appendix. 272
Figure 211: Cash flow diagram for the most probable set of operating profits 272
Figure 212: Three levels of predictions for the future operating profits 272
Figure 213: The triangular distribution which approximates more complicated types of distributions. Since values are uncertain, assigning values more accurate than high/medium/low is probably overly thorough 273
Figure 214: Triangle distributions for each year superimposed on the operating profits graph. 273
Figure 215: The first 10 (of many) operating profit simulations 273
Figure 216: Two hypothetical runs of a Monte Carlo simulation to illustrate the simulation method 275
Figure 217: Results of Monte Carlo simulation used to value the option of investing in R&D. 276
Figure 218: Flowchart of the simulation process 277
Figure 219: Present value of profit predictions at the decision point (year 1) 279
Figure 220: The decision tree approach can be shown to over value the option 279
Figure 221: Convergence of the Monte Carlo Simulation 280
Figure 222: Convergence of uncertainty in the Monte Carlo simulation 280
Figure 223: Three possible levels of unit costs for a PHM enabled circuit board 283
Figure 224: Pricing markup based on the uncertainty of the customers willingness to risk unplanned failures 283
Figure 225: Unit costs for the three scenarios 283
Figure 226: Possible levels of unit sales quantities 284
Figure 227: Estimated levels of engineering support for new PHM product 285

Figure 228: Three levels of predicted operating profits. A summary of this chart is presented as Figure 210 in the main body of the text 286
Figure 229: View of the micro-coil spring under an optical microscope at 20x magnification . 290
Figure 230: Schematic representation of the micro-coil spring 290
Figure 231: Micro-coil spring test vehicle 291
Figure 232: Representative drop test setup. For the data shown in this document the package was mounted face down on the drop table. The configuration shown in the picture allows the individual MCS to be monitored with high speed video during a drop event. 292
Figure 233: Relative displacement of package vs circuit board with MCS interconnects 293
Figure 234: Schematic of a Wheatstone bridge circuit 295
Figure 235: Wheatstone bridge with the resistance of the MCS daisy chain included as an arm in the bridge 296
Figure 236: Raw data recorded from Wheatstone bridge for test number one. 297
Figure 237: Calibration curve for Wheatstone bridge 298
Figure 238: Process flow of information through the various parts of the PHM system 299
Figure 239: Time history of five run to failure drop tests. The dotted red line represents the average resistance increase of each component at failure 300
Figure 240: Zoomed view of the resistance time history plots. The dotted red line represents the average resistance increase of each component at failure 301
Figure 241: Filter results for the data set from test one. The top plot is the filter track of the measured resistance. The bottom plot is the estimate of the ramp rate of change in resistance (ie: derivative). 306
Figure 242: Comparison of the numeric derivative and estimated derivative for test number one. Smooth derivatives are essential for accurately predicting remaining useful life and motivate the use of the Kalman filter 308
Figure 243: Remaining useful life predictions for the MCS 309
Figure 244: Alpha-lambda accuracy plot 310
Figure 245: Verification of code using the case of a damped harmonic system 311
Figure 246: Systematic error introduced by using an average failure threshold derived from the training set. 313
Figure 247: Process flow for leave one out cross validation discussed in this section and the comparison against traditional Weibull methods discussed in the next section 314
Figure 248: Residual between the estimated state and actual state demonstrates that the filter was tracking correctly for test number one 315
Figure 249: Residual between the estimate state and actual state for test #3. After drop 65 the filter divergence could have been used to warn that predicted RUL may be in error 316
Figure 250: Overlay of a few PHM predictions on top of the baseline weibull distribution..... 318

List of Tables

Table 1: Fatigue models commonly used for predicting reliability of electronics classified by approach. Reproduced from [Lee 2000] 21
Table 2: Types of Prognostics. Adapted from Saxena, A., Prognostics - The Science of Predictions, Conference of the PHM Society, 2010 28
Table 1: List of impedances for passive electrical components used in an AC Wheatstone bridge 41
Table 2: Typical values used in an AC Wheatstone bridge 42
Table 3: Component values for a low frequency AC bridge 46
Table 1: Inference with a perfect test and different false positive rates 80
Table 2: Inference with a perfect test and different initial beliefs in the system state 80
Table 8: Package Architectures for Test Board A 132
Table 9: Prognostic Distance (Test Board A, 127kHz) 144
Table 10: Prognostic Distance (Test Board A, 2.33MHz) 145
Table 11: Prognostic Distance (Test Board A, 6MHz) 145
Table 12: Package architectures on test board 169
Table 13: Anand's Constants for SAC305 173
Table 14: Undeformed geometry of solder ball 174
Table 15: Package architectures on test board 194
Table 16: Package Architectures used for Test Board B 211
Table 17: Discrete component values used in resistance spectroscopy AC Wheatstone bridge 216
Table 18: Parameter values used for damage model and particle filtering 223
Table 19: Package Architectures used for Test Board 234
Table 20: Full factorial test matrix used to study effect of simultaneous stresses and ageing effects 241
Table 21: Summary of fitting IPC failure data to a Weibull model 243
Table 22: Summary of critical resistances for each test 245
Table 23: Results of leave one out cross validation 256
Table 24: Micro-coil spring package details 290
Table 25: Drops to failure, and corresponding change in resistance 301
Table 27: Discrepancy between actual failure and failure threshold derived from training sets. 314
Table 28: Leave one out cross validation matrix. A cost metric of zero is a perfect score. 314

1 Introduction

Mission and safety critical electronics describes a system of electrical components whose correct operation are integral to the operation of a larger engineering system where an unplanned failure will result in loss of life or other unbearable consequences. As engineering systems continue to become increasingly electronic rich the need to assure the correct operation of mission and safety critical electronics increases in importance. This document will discuss methods for monitoring electronics for leading indicators of failure and predicting the expected future time of failure so that contingency plans can be made before a catastrophic failure occurs.

1.1 Mission and Safety Critical Electronics

Many electrical subsystems can be considered as mission and safety critical. Avionics systems require ultra-high reliability to fulfill critical roles in autonomous aircraft control and navigation, flight path prediction and tracking, and self-separation. Complex electrical power systems (EPS) which broadly comprise of energy generation, energy storage, power distribution, and power management, have a major impact on the operational availability, and reliability of electronic systems. Technology trends in the evolution of avionics systems point towards more electric aircraft and the prevalent use of power semi-conductor devices in future aircraft and space platforms.

Automotive applications demand many high reliability applications for electronics. In the power train engine controllers and the accompanying electronic throttle controls have received much scrutiny recently due to reports of unintended accelerations in automobiles. Autonomous vehicles, automated collision avoidance, air bags and anti-lock brakes require very high

reliability electronic systems to fulfill important safety applications. Automobiles are quickly becoming one of the most heavily mechanically/electrically coupled consumer products.

An increasing reliance on internet based storage and online applications requires very high reliability needs for computer servers. Many of the world's financial centers rely heavily on uninterrupted access to vast computer resources and large cities require extensive computing resources to maintain safety and order for millions of residents. As the world's population continues to shift to more urban environments their reliance on electronics for daily activities will increase substantially.

Long term human presence in space without the need for resupply is another typical example of mission and safety critical applications. Space travel tends to expose electronics to a variety of harsh environments, with severe consequences for unplanned failures and few opportunities for repair.

Due to the increasingly important role of electronics in modern engineering systems a number of traditional reliability approaches have been developed for electronics. A brief overview of traditional reliability methods will be discussed to frame a discussion about their drawbacks. Then new methods for overcoming the drawbacks of current methods will be introduced.

1.2 Traditional Reliability Approaches for Electronics

Traditional reliability approaches for electronics, also known as life prediction methodologies, use a combination of accelerated life testing and numerical simulation to describe the reliability of electronics. Even highly accelerated tests can be time consuming and

the accuracy of numerical simulations can vary widely due to the complicated nature of electronic assemblies and ever changing manufacturing procedures.

1.2.1 Accelerated Life Testing

Accelerated life testing is a method of using exaggerated environmental stresses to quickly damage an electronic assembly and gain information about its expected reliability. For purely mechanical stresses this typically involves thermal cycling, drop, and vibration testing. There are also a number of humidity, corrosion and electrical stresses that can be applied. Results from this type of testing are frequently communicated in the form of a Weibull failure distribution. To relate accelerated testing to actual usage conditions acceleration factors of varying quality may be available.

1.2.2 Numerical Simulation

In an effort to reduce the time and cost of physically testing electronics, a number of virtual qualification methods exist. There are some closed from tools that can quickly and roughly estimate reliability, but generally finite element simulation are used to analyze electrical components. Electronic assemblies are small in physical size, and incorporate many different constituent materials in a single product. Materials used in electronic components may have scale, temperature, and rate dependent properties which cause difficulties in successfully modeling the effect of applied stresses on components.

1.2.3 Hybrid Approaches

Hybrid approaches exist which combine experimental observations with simulation to predict reliability. For example the ability to measure strains in solder joints is very limited due to their small size (~0.5 mm height). As a solution to this problem strains near the solder joint on

the circuit board are measured, and then used as inputs to drive the boundary conditions for finite element simulations.

1.3 Shortcoming of Traditional Life Prediction Approaches

In a few closely controlled situations traditional reliability approaches can be used to successfully avoid unplanned failures. Unfortunately the way electronics are used in the field, particularly mission and safety critical electrical systems, the assumptions needed for traditional reliability methods to perform properly are often violated. In practice manufacturing variations and defects, simultaneous stress environments, unintended usages, and difficult to understand material properties undercut the effectiveness of traditional methods when an electronic system is mission or safety critical.

1.3.1 Drawback of Accelerated Life Testing

Typically accelerated life testing is performed on a batch of components from a prototyping line or a single manufacturing run. Accelerated thermo-mechanical testing for ball grid array (BGA), and chip resistor/capacitor components can take months to complete and significantly delays the timely launch of new products. Therefore there is strong monetary incentive to increase the speed of testing and draw conclusions from limited sample sizes. Using two accelerated tests to compare the relative reliability of two components can be effective provided a large enough sample size is used. It is much more challenging to acquire the data needed to determine acceleration factors that translate data from accelerated tests into estimates of field usage reliability. To generate these acceleration factors, you must make assumptions about the environment the electronics will be used in. If the actual usage differs from the assumed usage the quoted reliability from accelerated tests can be inaccurate. Another drawback of accelerated testing is the inability to handle manufacturing variations and defects. Test

assemblies from accelerated testing may not even be manufactured on the same line as the production parts. Variations in mechanical tolerances and properties can be difficult to identify during manufacturing because electrical testing (not mechanical) is predominantly used to validate the correct operation of circuit boards during assembly. If a small sample from each manufacturing lot is requisitioned for accelerated testing, months may pass between the start of a deviation in the manufacturing process and the identification of the problem due to the long running time of accelerated tests.

The most significant limitation on successful reliability life prediction is simultaneous usage conditions. Most accelerated testing involves imparting a single stress at a constant rate to a component. In the field, mission and safety critical electronics typically experience simultaneous stresses. For example automotive and aerospace systems typically experience widely varying temperature extremes and random vibrations simultaneously. In an effort to reduce the number of variables in an experiment accelerated testing usually does not stress components that may have interactions between two or more different stress types. As will be discussed later, a typical example is the reliability of a solder interconnect subjected to a drop event can be drastically reduced by prior isothermal ageing.

1.3.2 Drawbacks of Numerical Reliability Prediction Methods

Numerical simulations suffer from a number of drawbacks when trying to predict reliability in mission and safety critical electronics. Most importantly the simulations are difficult to directly validate due to the small size of the critical parts of an electronic assembly. Second order validations are used because as of yet no one has figured out how to experimentally observe strains in the tiny critical locations that are mechanically weak. An example of a second order validation is to use a correlation between a simulation and experimentally observed failure

data when in fact we are truly concerned with un-observable stresses inside the component. The danger of *correlation without causation* is much higher with a second order correlation. Additionally simulations must assume ideal geometries and pristine material properties. Due to normal variations in manufacturing, geometries may not be adequately modeled. Manufacturing defects are seldom considered. The non-linear, scale, temperature, and time dependent properties of materials used in electronics makes building a successful simulation challenging. In situations where stresses are applied in a dynamic method such as drop and shock the computational expense of numerical analysis of can be very high.

1.3.3 Drawbacks of Hybrid Reliability Approaches

Hybrid reliability approaches suffer from a combination of experimental and numerical challenges discussed in the previous two subsections. Again the single stress nature of accelerated tests may limit the applicability of hybrid approaches to mission and safety critical electronics. Many times hybrid approaches rely on strain values extracted from numerical simulations which are nearly impossible to validate by experimental methods. Interacting factors that are seen in fielded products, but not easily duplicated in a laboratory, can also limit these approaches if reliability predictions were developed from assemblies undergoing a single stress test.

1.4 Prognostic Health Management

The limitations of traditional reliability methods when applied to mission and safety critical electronics in realistic fielded environments has motivated the development of prognostics and health management methodologies (PHM) for electronics. Named after an analogy from the medical field where patients are given a prognosis for their survival time based on medical indicators specific to the patient, prognostics in an engineering discipline which

attempts to predict the remaining useful life (RUL) of a physical system. Similar to the medical profession, leading indicators of failure are used to help quantify the damage state of the system being monitored. The development of PHM has tracked closely with the evolving methodology of maintenance. Maintenance methodologies have evolved from reactionary to performing time based and condition based maintenance. In practice as the reliability of engineering systems has increased a larger number of applications are either becoming, or desire to be, mission and safety critical.

Assuming traditional reliability methods can be applied successfully, which may not always be true; you are encountered with the situation where you must choose the acceptable probability of failure in a group of systems. This is represented schematically in Figure 1. The area under a predicted failure distribution curve is defined as unity. Therefore an N% allowable probability of failure can be identified as a region under the curve with an area of N%. The edge of this area defines the N% failure time. Traditional maintenance approaches would schedule maintenance or replacement of a component when its usage exceeds the N% threshold. In practice high reliability systems may be thresholded at a maximum probability of 1, 0.1, or even 0.01% failure. This results in two undesirable situations. The first disadvantage is there will always be a small probability of an unplanned failure. By definition if the system is mission and safety critical this is unbearable. Furthermore (100 – N)% of components will be replaced before they are truly worn out which can make maintaining a mission or safety critical system very expensive. If the traditional reliability method that was used to derive the predicted failure distribution is flawed in any of the ways described earlier, the N% probability of failure may not accurately represent the true risk of failure. Most importantly traditional reliability methods use large population statistics to describe reliability as an average in a large fleet of identical

products. Traditional methods provide no means for predicting the reliability of a single component given its truly unique history of past usage and infinite number of possible future usage conditions.

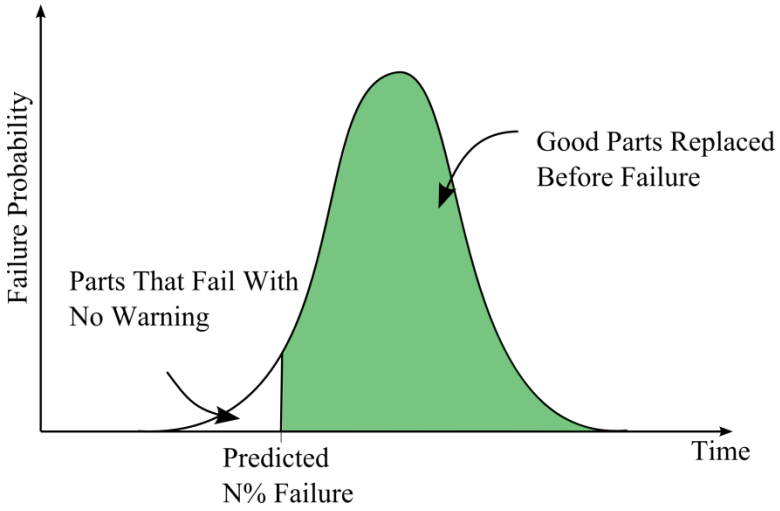


Figure 1: Hypothetical failure distribution highlighting some drawbacks of traditional reliability methods

PHM methodologies are a reaction to traditional methods. PHM uses advances in sensor technology and failure models to monitor a component as it is in use. Predictions of remaining useful life can be made based on the true state of the system so that contingency plans can be made before failure occurs. If a manufacturing defect has escaped detection and is installed on a fielded system PHM can provide warning that an earlier than traditionally expected failure may occur. If the system is used in a manner or environment that it was not designed for, PHM can provide warning of an impending failure because actual damage to the system is monitored. Damage accrued during shipping, handling, and maintaining the system can also be captured using PHM. There are many aspects to PHM. Damage isolation and detection can be used to identify that the system has accrued damage. Failure mode classification methods are used to determine what failure mode is most dominant. This work will focus on the prognosis aspect of PHM including the development of sensors to monitor damage and models for predicting

impending failures. Much of this work will ultimately focus on determining the remaining useful life of a system given a measurement of the current state.

1.5 Risk Based Decision Making

"Knowing when a system will fail is not nearly as valuable as knowing when to take an action." – Steven Engle

Risk based decision making is a methodology that embodies how to make prognostic predictions of failure useful for the end user of a system. Using a process established in [Engle 2009], statistically defendable methods for using predictions and taking action based on those predictions is established. Using the same remaining useful life prediction in different ways, risk based decision making can be structured to optimize, safety, cost, availability, and logistics. Many of the techniques used in this document were chosen because they ultimately facilitate the use of risk based decision making.

1.6 Resistance Spectroscopy

Resistance spectroscopy is a method of making resistance measurements on very small specimens, like those found in electronics. The denotation spectroscopy implies that the measurement has a magnitude and phase component, like a sine wave, that is used to characterize the system being measured. The ability of the resistance spectroscopy measurement approach to monitor small changes in resistance in a very noisy environment makes it an ideal sensor for non-destructive evaluation of damage in fielded electronics. The entirety of this work is made possible by the measurement capability of the resistance spectroscopy method. Earlier when PHM was described as leveraging advances in sensor technology to provide unique capability beyond traditional methods, the resistance spectroscopy method is the advancement

that makes this work feasible. The choice of resistance spectroscopy as a practical sensor for PHM of electronics is crucial because it can be easily implemented in a cost effective manner on circuit boards using the same technology that is used to manufacture the rest of the circuit board. The prevalence of easy access to the electrical signals and proximity to failure locations on a circuit board makes it ideal for prognostics when compared to other more mechanical measurement methods.

1.7 Overview of Electronic Packaging

Much attention has been given to Moore's Law [Moore 1965], which roughly states that based on prior experience the number of transistors that can be cheaply built into a single chip will double every 24 months. Assuming that a model for the number of transistors per chip takes the form of

$$\mathrm{N(t)} = \mathrm{C}2^{\frac{\mathrm{t}}{d}} \tag{1}$$

Where N is the number of transistors per chip, C is a constant, t is time in years, and d is the doubling period in years. The regression problem can be made linear using a base two logarithm.

$$\log_2 \mathrm{N(t)} = \log_2 C + \log_2 2^{\frac{\mathrm{t}}{d}} = \log_2 C + \frac{\mathrm{t}}{d} \tag{2}$$

Based on the data presented in Figure 2, the realized transistors per chip still closely fits to the model proposed by Moore in 1965. The data presented in Figure 2 indicates that the number of transistors have doubled every 1.991 years between 1971 and 2012.

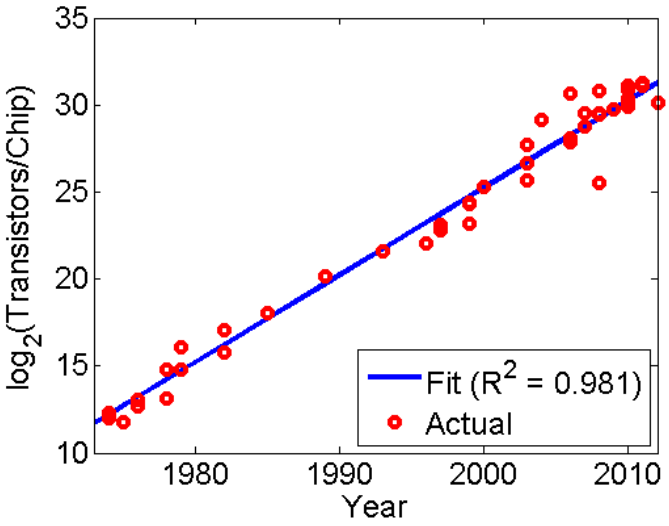


Figure 2: Transistors per Chip Realized in Practice

The unintended ramification of doubling the number of transistors on a chip every 24 months is the need to connect an increasing number of interconnects between the micro scale of the semiconductor, and the macro scale of the physical world where electronics interact with humans. The process of protecting the fragile integrated circuit from the environment in general is known as the field of electronic packaging. Figure 3 shows a number of the different levels of packaging involved in a complex electrical system

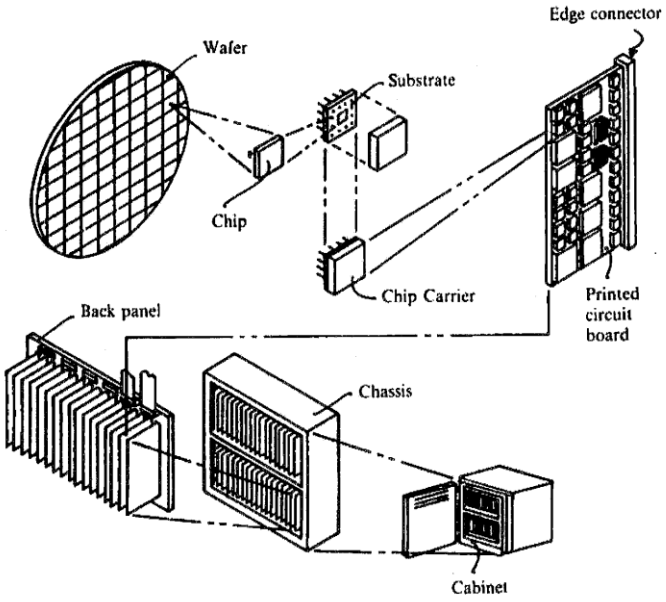


Figure 3: Various levels of electronic packaging [Dally 2008]

First and second level interconnects, or the connections between the chip and the printed circuit board bridge the interface between the micro level and macro level and often are challenging to make and reliably maintain. On a typical CPU mother board there are thousands of second level interconnects, and a single interconnect failure can render the entire computer useless. As the complexity of integrated circuits increases, the number of required second level interconnects increase. Figure 4 highlights the space advantages of area array packaging as compared to simply connecting around the periphery of a chip, and Figure 5 shows a common format for high I/O packaging based on this concept called a ball grid array (BGA).

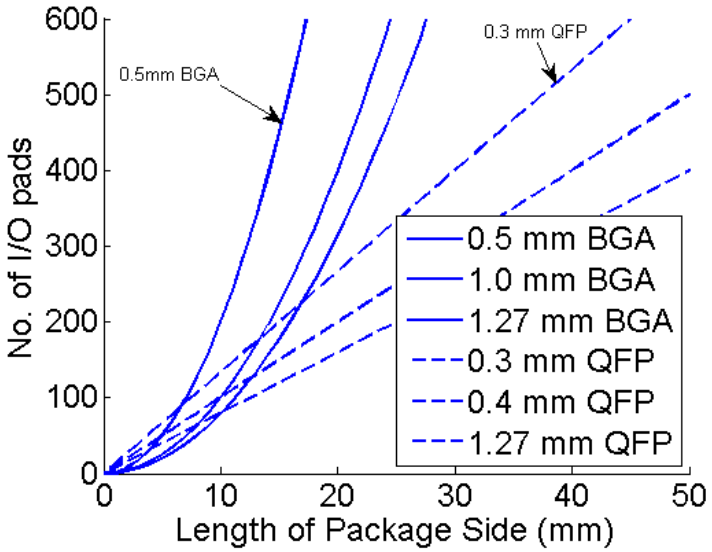


Figure 4: Interconnects per length for different packaging architectures [Dally 2008]

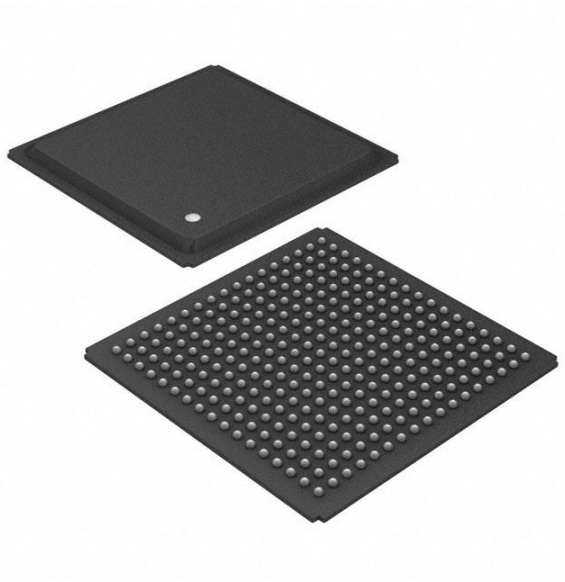


Figure 5: Widely used format for electronic packaging, an array of solder bumps or solder joints. Top view (upper), bottom view (lower). The bottom side connects to the printed circuit board. The integrated circuit is enclosed and protected by the black mold compound.

Much of the work in this document focuses on monitoring and predicting failures in grid array packaging. Figure 6 shows an idealized cross section of two common grid array packages. The flip chip is a more modern technology where the integrated circuit is connected directly to the circuit board using solder joints. The BGA package is a slightly taller configuration where an interposer circuit board connects via wire bonds to the integrated circuit and then solder joints connect the interposer board to the printed circuit board.

Figure 6: Cross section of a flip chip package (left) and a BGA package (right)

A number of environmental stresses such as drop/shock, vibration, temperature, and many others can cause fatigue cracks to propagate through solder joints as depicted in Figure 7

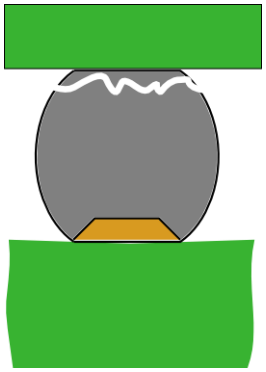


Figure 7: Depiction of a cracked solder joint as the result of applied environmental stresses. A crack in a solder joint causes an open circuit and failure of the electrical system

At the time this document was created flip chip and ball grid array packaging was in widespread use. New architectures designed to further increase the density of interconnects by utilizing all three physical dimensions had begun to be used in higher end electronics. Undoubtedly the march of progress will continue in the electronics industry, and each new iteration of integrated circuits will be susceptible to the mechanical limitations and reliability concerns of the electronic packaging that protects the integrated circuit. In many consumer applications that utilize cutting edge technology a certain level of unexpected failures are considered acceptable. In mission and safety critical electronics the consequences of unplanned failures are unbearable and motivate the use of new techniques to monitor and warn of impending failure. This work specifically attempts to create methods of monitoring solder joints before a fatigue crack propagates through the joint and predicting when the system will be rendered unusable so mitigating action can be planned.

2 Literature Review

Traditional reliability methods for electronics have been covered in extensive detail [Minges 1989, Pecht 1991, Lall 1996, Blackwell 2000, Harper 2005, Dally 2008]. A brief overview of reliability methods will be made to substantiate claims made in the introduction of this document. A discussion of traditional reliability methods highlights drawbacks of the current approaches when applied to mission and safety critical electronics. The foundation for improvements to traditional methods via sensor advancements will be reviewed in the resistance spectroscopy section. Prognostic health management in general and prognostics for electronics in particular will be reviewed to provide a context for the contributions of the current work.

2.1 Traditional Reliability Methods

Traditional reliability methods focus on gathering large population failure data through accelerated life testing, virtually qualifying electronic assemblies through simulation and hybrid approaches that combine experimental data with simulation.

2.1.1 Accelerated life testing

Accelerated life tests (ALT) are a type of tests that is designed to quickly fail a component in a manner that is representative of how the product might fail during actual use [Suhir 2002]. This work focuses on direct mechanical stresses that can be imparted on electronics by the environment electronics are typically used in. Examples of environmental stresses include thermo-mechanical stress due to mismatch in coefficients of thermal expansion as ambient temperatures change, drop loadings caused by impacts during use and accidental falls during transportation and maintenance, and vibration loadings caused by transportation and usage. ALTs may be combined, for example simultaneous vibration and elevated temperatures.

Other accelerated tests that will not be discussed include stresses imparted by corrosion, humidity, chemical and electrical stresses. The US Department of Defense (DoD) military test standard MIL-STD-810 list a large number of different accelerated tests that may be used for a variety of different products including, but not limited to electronics. Accelerated tests standards for surface mounted electronics common in mission and safety critical electrical products have been adopted by the electronics industry for thermomechanical stresses [IPC-SM-785], drop [JESD22-B111], and to a lesser extent vibration [JESD22-B103B]. Sometimes vibration profiles outlined in MIL-STD-810 are used or modified for vibration testing of surface mount electronics.

In addition to describing how to run an accelerated test, standards provide insight into methods of monitoring electronics for failure during the accelerated test. Electronics subjected to highly dynamic stresses such as drops, shock, and vibrations must be monitored using high sampling rate equipment to catch brief intermittent failure events. Standard practice is to assemble a component under test so that each electrical interconnect is connected in a single series circuit, or a continuity daisy chain. If any of the electrical interconnects fail mechanically, the resistance of the daisy chain becomes infinity and can be easily identified. Due to the dynamic nature of many of the ALTs broken daisy chains can appear to be not broken when the test is not running. Therefore the severity, number, and duration of open events that defines a failure are dependent on the test standard and must be monitored continuously.

Accelerated life test data is often communicated in the form of Weibull statistics. The use of Weibull statistics is partly based on science and partly based on historical reasons [Hallian 1993]. The Weibull failure distribution can be configured to widely different shapes and covers many practically observed failure phenomena, specifically components that do not have a constant failure rate. To produce accurate statistical estimators, particularly for lower percentiles,

a large sample size is required [De Souza 1995]. Due to time to market and budget pressures, in practice the number of samples is usually less than optimal. There are at least three standard forms for the Weibull distribution [Hamada 2008], one of which can be stated as

$$f(x) = \alpha\eta x^{\alpha-1}\exp(-\eta x^{\alpha}) \quad (3)$$

Where f is the probability density function (PDF) of failure at time x, α is the shape parameter and η is the scale parameter. Different methods exist for fitting failure data acquired through ALT or from field data to the Weibull model. There is a graphical method, maximum likelihood estimate (MLE), and Bayesian approaches to estimating the Weibull parameters [Hamada 2008]. The MLE method is used most commonly and can be found in the WeibullSmith [Fulton 2002] and Matlab [Mathworks 2012] software tools. To quantify the effect of sample size on the effectiveness of the MLE estimate of the shape and scale parameter a simulation was created. First a ground truth distribution was created with values of α and η as 1.5 and 0.7 respectively (Figure 8). Then a limited number of random values were sampled from the ground truth distribution. The MLE approach was used to estimate the Weibull parameters from the limited sampled data. The process of drawing samples and estimating parameters is considered one trial. Ten trials were performed at each of the different sample sizes. Mean and standard deviations for each group of ten trials at a specific sample size was calculated.

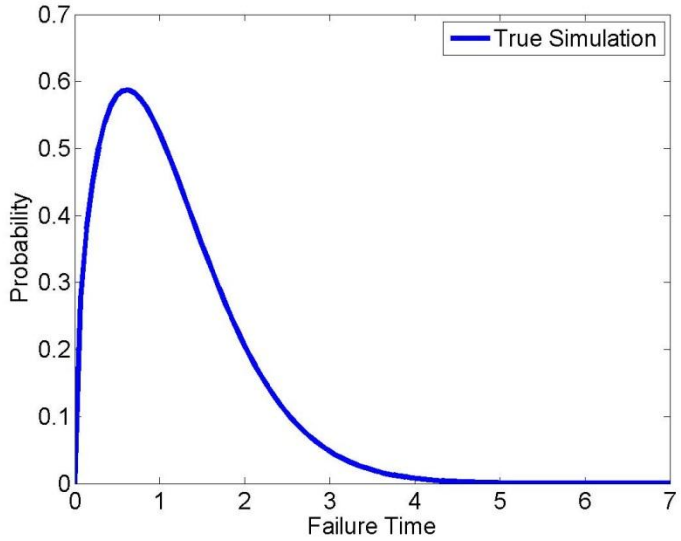


Figure 8: Ground truth PDF used for the Weibull accuracy simulation

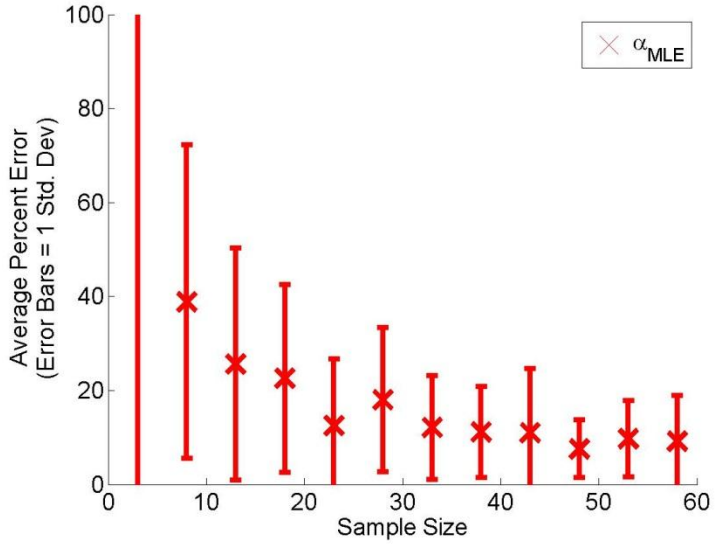


Figure 9: Error in MLE parameter estimation for different sample sizes compared to the ground truth simulation

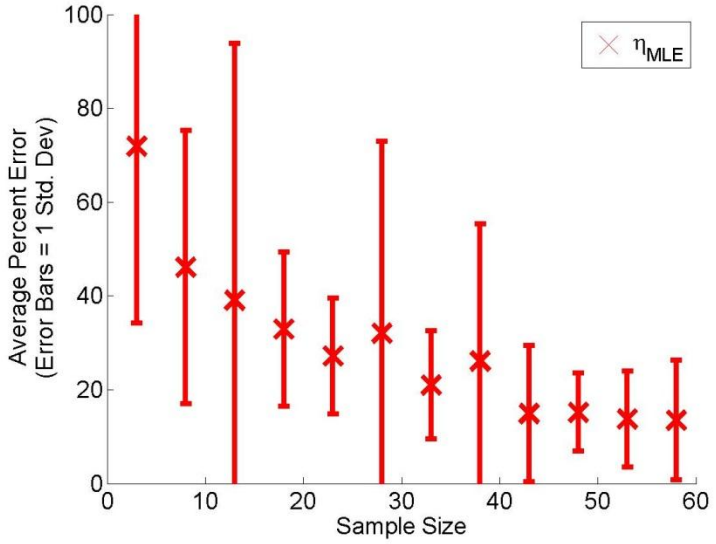


Figure 10: Error in MLE parameter estimation for different sample sizes compared to the ground truth simulation

The simulations highlight the errors that may occur when estimating the shape (Figure 9) and scale (Figure 10) parameters using the commonly used MLE method with even moderate sample sizes. The resulting estimates may not accurately provide sufficient resolution to predict low percentile thresholds needed for mission and safety critical electronics.

2.1.2 Numerical Simulation

A limited number of closed form approaches exist for life estimation in electronics. Closed form methods for thermal cycling [Clech 1993], and vibration [Stienberg 2000] environments have been proposed. Most life prediction approaches use some form of finite element modeling. In general finite element approaches involve choosing a constitutive model to describe the behavior of materials used in electronics, assumptions about geometry and boundary conditions are made and representative models are created. Commercial software calculates stress-strain values for the model. Stress and strain values from the analysis are used as inputs into one of the many solder fatigue models to predict the expected time of failure for a component. Simulations concerning reliability due to thermomechanical, drop and vibration

loadings focus on solder joint reliability as the mechanically weak component. Failures in other components may not be studied and can lead to unexpected failures. Constitutive models used for modeling solder must account for time and temperature dependent properties [Hongtao 2006, 2007, Yeo 2006, Bhate 2008] to achieve reasonable accuracy in life prediction. Different approaches to modeling geometry have been used. Smeared models [Pitarresi 1991], global/local submodeling [Lall 2007f], and non-traditional problem formulations [Lall 2010g,h] have facilitated the modeling of electronics.

Lee provided a concise review of solder joint fatigue models used in electronic packaging [Lee 2000]. A summary of the fatigue model names and approach classifications reproduced from the cited work is shown in Table 1. Each model tends to be appropriate for a specific loading condition, geometry, or material type (e.g.: SnPb or SAC alloy solders). Strain based approaches assume that an applied strain causes a stress in the solder, as is common in thermal cycling applications. The author further divides the strain based criteria's into models that account for plastic strain or creep strain dominated effects. Energy based approaches utilize information from a stress strain hysteresis loop to predict fatigue life. Damage based approaches utilize fracture mechanics approaches to predict fatigue failure. The author provides an insightful warning that fatigue life and solder joint life are not necessarily defined as the same quantity because a solder joint with a growing fatigue crack can still maintain electrical continuity in some cases. Famously a bug in the commercial finite element software ANSYS necessitated a correction to published model constants in the frequently used Darveaux model [Darveaux 2002]. The same model was updated with new material constants when lead free solders became widely used [Lall 2009].

Table 1: Fatigue models commonly used for predicting reliability of electronics classified by approach. Reproduced from [Lee 2000]

Fatigue model	Strain		Energy	Damage	Other
	Plastic	Creep			
Coffin-Manson	X				
Total strain	X				
Soloman	X				
Engelmaier	X				
Miner	X	X			
Knecht and Fox		X			
Syed		X	X		
Akay			X		
Liang			X		
Heinrich			X		
Pan			X		
Darveaux			X	X	
Stolkarts				X	
Norris and Landzberg					X

Historically solder for electronics applications was an eutectic formulation of tin-lead solder commonly abbreviated as Sn63Pb37. Due to directive 2002/95/EC , restriction of the use of certain hazardous substances in electrical and electronic equipment (RoHS) by the European

parliament and council, usage of solders containing lead have been drastically reduced. Starting in 2006 solder formulations based mostly on tin and silver known as SAC alloys have been predominantly used. Unfortunately for reliability engineers there are a wide number of different chemical compositions trying to obtain market dominance in the lucrative electronics manufacturing industry. Slightly different solder compositions can have drastically different material properties that evolve with time and temperature [Hongtao 2009]. Obtaining adequate material properties to build finite element models that can accurately model deformations in lead free solder joints can require many difficult to perform tests beyond the classic room temperature quasi-static pull test [Mustafa 2011, Lall 2011d].

The electronics manufacturing industry is dominated by consumer electronics due to the large volumes of cell phones, computers, gaming systems, tablets, and related products that are manufactured each year. The entire group of manufactures who specialize in mission and safety critical electronics has negligible buying power compared to the consumer electronics sector. The unintended consequence of the RoHS initiative is the disappearance of SnPb electronic products from the global supply chain. Each year the economics of maintaining obsolete equipment for a niche market of mission and safety critical electronics customers becomes less advantageous so a change to lead free alloys is being forced on many consumers. In response to reliability concerns caused by the lead free transition, a coalition of stakeholders including NASA, DoD, DoE, Army, Navy, Air Force, defense and space contractors are studying risk mitigation strategies. The NASA JCAA/JGPP project, and the Pb-free Electronics Risk Mitigation consortium (PERM, aka: Pb-free Manhattan project) are examples of the large bodies of work being amassed on the topic of the lead free solder transition for mission and safety critical electronics.

2.2 Resistance Spectroscopy

Resistance spectroscopy is a measurement technique that is used extensively in later parts of this document. Originally developed for electronic packaging [Constable 1994, 1995] and then identified as a reliability tool applicable to ALTs [Butler 2000]. Implementations of the resistance spectroscopy technique using commercial off the shelf hardware and the Labview programing language [Batra 2003] show an economical method for implementing the technique. The underlying concept of the resistance spectroscopy technique has historically been used for a number of applications [Temple 1973] in the physical sciences. A full section is dedicated to the specifics of this technique later in the document.

2.3 Prognostic Health Management

Prognostic health management is a paradigm shift away from reactive fear of failure approaches associated with maintaining high-reliability critical electronics systems. The benefits of monitoring a component for early indicators of failure, prognosticating future failure, and then taking action before a catastrophic failure occurs have many applications [Luna 2009, Grubic 2009]. Monitoring the health of individual components enables improvements in safety [Downes 2007], availability [Swanson 2001], and cost control of high reliability systems [Jarrell 2002]. Traditional reliability methods, such as Weibull analysis, provide failure time estimates for a large population of components. Unfortunately the failure of a single electronic component would be the most useful information, since single components - not average components - cause system wide failures. Knowing the health of single components provides substantially better information for making decisions [Engle 2009] that affect safety, availability, and cost of high reliability systems. Advances in sensor technology and failure analysis have catalyzed a broadening of application scope for prognostication systems to include large electromechanical

systems such as aircraft, helicopters, ships, power plants, and many industrial operations. Current PHM application areas include, fatigue crack damage in mechanical structures such as those in aircraft [Munns 2000], surface ships [Baldwin 2002], civil infrastructure [Chang 2003], railway structures [Barke 2005] and power plants [Jarrell 2002]. The spectrum of reliability and health monitoring approaches for electronics has traditional methods at one end of the spectrum, hybrid reliability approaches, such as fuses and canaries, in the middle of the spectrum, and prognostic health monitoring approaches at the far end of the spectrum.

Life Consumption Monitoring is a methodology that at times has been described as prognostics [Ramakrishnan 2003, Mishra 2002, 2004], but the author feels that it is more accurately described as a usage monitoring maintenance approach [Hopp 2011]. Life consumption monitoring uses observations of usage conditions as inputs into finite element models to perform a customized approach to traditional numerical reliability approaches. Because environmental conditions, and not the physical state of the system, is being monitored there is no method for detecting manufacturing defects, material variations, or other unplanned realities that are not easily captured by a priori computer modeling.

Reliability canaries represent another method that falls on the spectrum of reliability methods between traditional methods and prognostics. Named after the song birds that coal miners would carry into mines to provide advanced warning that toxic gases were accumulating in the working areas. In the mines, the small bird would die when exposed to concentrations of toxic gases that were non-lethal, but potentially dangerous to a human. On a circuit board, a canary is a usually a small nonfunctional component that will fail prior to the expected failure time of the functional components [Vichare 2006]. Again the problem when applied to mission

and safety critical applications is the lack of insight into the functional components and no assurances that the manufacturing defects or anomalies will be discovered.

For a long duration space mission, prognostic technology would be required to be integrated (embedded) into the systems being monitored, consume minimal resources, and provide sufficient advanced warning of failure to allow contingency plans to be formulated. Some approaches that may meet these goals in other application domains and can be loosely characterized into structural applications or energy conversion applications. Techniques related to structural monitoring include: bridges [Ko 2005], aircraft [Muns 2000], railways [Barke 2005], and nuclear power plants [Jarrell 2002, Coble 2010], Alternatively other authors have investigated monitoring power delivery and conversion system components: batteries [Goebel 2008], generators [Swanson 2001], planetary carrier plates [Orchard 2007, 2009], bearings [Bechhoefer 2011], and split torque gearboxes [He 2010]. Grid array components discussed in this paper are part of the subset of electrical energy distribution and conversion. Different electrical components tend to have different failure modes even though they are found near grid array components on circuit boards. For example: electrolytic capacitors [Kulkarni 2010], gate bipolar transistors [Brown 2011], and MOSFETS [Saha 2011b].It is important to note that while this paper mainly focuses on monitoring and prediction, the algorithm outputs have been specifically designed to dovetail with the management aspect of PHM. Risk based decision making methods [Engel 2009, Lall 2010b], allow the outputs from a prognostic prediction to result in actionable decisions for an end user.

Traditional reliability methods have implicit shortcomings that are often unavoidable. Critics of traditional reliability methods often point to the need to capture high-resolution environmental data on prior usage [Jiang 2009], discontinuity between accelerated test and field-

usage conditions, sensitivity to manufacturing defects [Elerath 2004], and the reliance on large population statistics when single components statistics are most relevant to the repair of a specific system [Hamada 2008]. Reactive maintenance, where parts are replaced as they fail, often is not an option for high-reliability electronic systems. Thus, electronic components or assemblies may often be replaced at a very low probability of failure - typically characterized by a p value of one-percent or lower depending on application. The disadvantage of this approach is that most components are replaced well before they wear out, leading to excessive and often avoidable support costs. Further, in many applications it may not always be possible to collect run-to-failure test data. Fuses and canaries may provide discontinuous data on damage initiation and progression. Prognostic health management, based on leading indicators of failure eliminates the need for insight into prior usage, reliance on large population statistics, and insight into variance in manufacturing conditions. The use of leading indicators of failure provides the ability to address time varying or evolving usage conditions. Leading indicators can be used to identify manufacturing defects often characterized by assemblies which degrade at unusually rapid rates. Furthermore, leading indicators based PHM is component specific and is not affected by uncertainties associated with conclusions drawn from large population statistics. Previously damage initiation, damage progression, and residual life in the pre-failure space has been correlated with micro-structural damage based proxies, feature vectors based on time, spectral and joint time-frequency characteristics of electronics [Lall2004a-d, 2005a-b, 2006a-f, 2007a-e, 2008a-f]. Commonly prognostic algorithms which incorporate a forecasting component are implemented though recursive algorithms such as a Kalman filter or Particle filter.

Kalman filtering is a recursive algorithm that estimates the true state of a system based on noisy measurements [Kalman 1960, Zarchan 2000]. Previously, the Kalman Filter has been used

for navigation [Bar-Shalom 2001], economic forecasting [Solomou 1998], and online system identification [Banyasz 1992]. Typical navigation examples include tracking [Herring 1974], ground navigation [Bevly 2007], altitude and heading reference [Hayward 1997], auto pilots [Gueler 1989], dynamic positioning [Balchen 1980], GPS/INS/IMU guidance [Kim 2003]. Application domains include GPS, missiles, satellites, aircraft, air traffic control, and ships. The ability of a Kalman filter to smooth noisy data measurements is utilized in gyros, accelerometers, radars, and odometers. Prognostication of failure using Kalman filtering has been demonstrated in steel bands and aircraft power generators [Batzel 2009, Swanson 2000, 2001].

The particle filter has been used for projectile tracking [Arulampalam 2002, Gustafsson 2002, Ristic 2004], robot localization [Verma 2004], fatigue crack growth estimation [Cadini 2009], and prognostics for helicopters [Orchard 2007,2009], batteries [Saha 2009a,b], and pneumatic valves [Daigle 2009, 2010]. In general, particle filtering is classified as a sequential Monte Carlo technique [Doucet 2001]. Particle filtering may also be known as sequential importance sampling, bootstrap filtering, or Monte Carlo filtering. The benefits of using the particle filter are a fully encompassed methodology for tracking nonlinear or non-Gaussian damage variables with seamless incorporation of uncertainty management [Saha 2008] into the failure prediction algorithms. The use of particle filter in conjunction with resistance spectroscopy and phase-sensitive detection for electronics PHM is new. Methods specifically designed for quantifying the performance of prognostic algorithms [Saxena 2008a,b, 2009a,b] can be used to report results in a concise manner. A full discussion of prognostic performance metrics is included later in the document.

Fault mode identification and classification is a paradigm that initially was developed for diagnostics [Padalkar 1991, Bernieri 1994, Isermann 1997], and was later employed in the pre-

failure space [Lall 2010e,f] for prognostics. Fault mode identification is useful for detecting changes in a system as it begins to fail or have unexpected behavior. Fault mode classification attempts to predict the dominant failure mode, after an anomaly has been detected in a fault mode identification scheme, when there are many possible failure modes that can occur. Prognostics can be classified into three stages [Schwabacker 2007]. Typically stage one is the fault identification stage, stage two is the fault classification stage, and stage three is the prognosis stage. Stage three may alternatively be described as the stage where remaining useful life is estimated. The prediction stage is acknowledged as the least developed stage in prognostics, and this document will focus exclusively on innovations in stage three for predicting failure in electronics. Table 2 highlights different reliability methods that contain components of the prognostics methodology, or have been described as prognostics by other authors. The table is presented in an attempt to disambiguate contradictory terminology used in the literature. Condition based prognostics are taken as the desired application for prognostics discussed in this document.

Table 2: Types of Prognostics. Adapted from Saxena, A., Prognostics - The Science of Predictions, Conference of the PHM Society, 2010

Name	Description
Reliability Based Data	Uses population based statistical models and historical time to failure records. Estimates average life of an average component under typical usage conditions. Examples: Weibull analysis, acceleration factors.
Stress Based	Uses knowledge of failure modes and assumptions about environmental conditions to predict average life of a component under a specified usage condition. Examples: Finite element analysis, Coffin-Manson/ Darveaux models, Proportional hazard models.
Condition Based	Uses real time information from sensors to measure or infer component degradation. May utilize data driven, model based, or hybrid approaches. Estimates remaining life for a single

	component under assumed future usage conditions. Examples: Cumulative damage model, filtering and state estimation.

The methodologies of Verification and Validation (V&V), can be used to rigorously demonstrate the capability of a system to show that it was built correctly and performs as intended. V&V is a popular technique in system engineering [Kapurch 2010] and computational mechanics [ASME 2006]. It is easiest to understand V&V from a systems engineering standpoint, but the methodologies used in prognostics are probably closer related to computational mechanics since they rely heavily on algorithms and differential equations. From a systems engineering perspective, V&V can be loosely stated as

- Verification: Did you build your product correctly to the specifications of the technical drawing
- Validation: When you use your product, does it perform as intended

It was identified that V&V would be both necessary and difficult to achieve for PHM algorithms [Jacklin 2004, 2005, Roemer 2005, Hao 2011], but there appears to be no advancements to the state of the art specifically for PHM. Outside of the PHM community, a number of seemingly unrelated researchers are developing V&V techniques that indirectly apply to prognostics. Physics based models are validated and verified using methods described by [Hills 2005] for heart transfer and computational mechanics models. In [Sornette 2008] a general approach to model validation is described, using earthquake, solar diffusion, and fluid dynamics models and illustrative examples. PHM algorithms commonly contain data driven components of machine learning algorithms, such as neural networks [Lall 2010e], that may not be easy to validate because they are not a closed form solution. Guidance for validating neural network

controllers is provided by [Pullum 2007]. At this time the author is not aware of a demonstrated comprehensive method of V&V for prognostics.

As a historical note, prognostics borrows its name from the field of survival analysis. Survival analysis is the study of predicting the time to an event, where the event is typically death. Survival analysis is a common name for time to event analysis in the health sciences. Equivalent analysis are performed as reliability analysis in engineering fields [Bell 2004], duration analysis in economic fields [Klein 1991] and under varying names in fields as varied as demographics, astronomy and insurance. Since discussing death in humans, as opposed to failure in electronics, carries a certain stigma most health related analysis are performed needing to ascertain survival not failure, and hence the frequent references to prognostics in medical fields of study.

Actuarial tables (life expectancy tables) are a method based solely on historical death records, and have historically been the corner stone of the insurance industry [Dawson 1898, Shang 2006]. Medical studies are performed to evaluate the effectiveness of a treatment for a specific disease. A technique known as the Kaplan-Meier curve [Kleinbaum 2005] extends the data driven actuarial method to compare a treatment group and placebo group. The Kaplan-Meier curve also allows the evaluation of a medical treatments effect on survivability as a function of time and can accommodate censored data. Logistic regression [Klienbaum 2002] is a type of regression that is applicable to binary variables (life or death), and commonly used in survivability analysis. Cox proportional hazard models [Bell 2004], are a modern approach to survivability analysis. Cox models have the capability to compensate survival models for competing variables such as sex or ethnicity [Moriguchi 1993]. In [Klienbaum 2005] it is shown that Weibull reliability models can be algebraically manipulated to be equivalent to proportional

hazard models used in survivability analysis. The same issues that restrict the use of traditional reliability data based approaches for prognostics also restrict the use of related survivability analysis methods.

Weather forecasting is a discipline that would seem to have similarities to PHM on many levels, only applied to a different application area. While some of the verification methods for forecasting [Roebber 1996] have been borrowed for prognostics, direct insight is difficult to obtain due to the highly application specific nature of weather forecasting. In weather forecasting a variety of measurements such as airport weather stations, radar [Whitton 1997], satellites [Kuciauskas 2010], and ocean floats [Roemmich 2009] are combined with weather models. Knowledge of the interface between the atmosphere and the planet is known to be critical for prediction accuracy, but is difficult to monitor with sufficient resolution. A combination of statistical models and dynamic weather models (mathematical models based on differential equations) are used to forecast future weather. Dynamic weather models are based on differential equations that govern the physics of the earth's atmosphere [Al-Yahyai 2010] and can be grouped into global weather models, and local area models. Global and local models represent a tradeoff between efficiency, accuracy, and resolution. Traditionally the accuracy of local area models is better than global models, but accuracy dramatically reduces as a result of influences outside the modeled area after 48 hours. Famously Lorenz correctly predicted that due to the chaotic nature of the fluid dynamics involved in weather forecasting, predicting the weather accurately beyond 14 days will be beyond the state of the art [Lorenz 1965] for the foreseeable future. Lorenz's prediction highlights the nature of dynamic weather models as simply an initial value problem. State of the art global weather models [Ruston 2012] have a grid spacing that is typically spaced 25 km apart up to an elevation of 65 km. Typically this results in a model

meshed with about $21x10^6$ nodes that will execute on a 12 teraflops super computer in 12 hours. By comparison local weather models have nodes at 10 km increments and can run on a modern personal computer. Dynamic weather models did not surpass the accuracy of statistical models for the five day forecast until the early 1990's [Rapport 2009]. Data assimilation, the fusion of model and observations, relies on variational calculus to minimize error between model and observation states in three spatial dimensions and one time dimension [Navon 2009]. A technique known as an ensemble Kalman filter, which avoids inverting matrices of observations that may have a million entries, was specially devised for data assimilation in weather forecasting models [Evensen 1994, 2003].

3 Resistance Spectroscopy

The resistance spectroscopy (RS) technique provides the experimental foundation for the work presented in this document. The method is quite elegant when its application is distilled into its components. Essentially the technique is a very narrow band pass filter, but is implemented in an unexpected manner compared to traditional noise filters. The ability of the filter to reject noise at frequencies not related to the signal makes the technique very robust. Lastly the applicability of the technique using different test conditions is verified to match with theory.

3.1 Direct Current Wheatstone Bridge

A Wheatstone bridge is a collection of resistors organized in a manner that facilitates the measurement of changes in resistance [Wheeler 2004]. The bridge is widely used in mechanical strain gages [Dally 2005]. A diagram of a typical DC Wheatstone bridge can be seen in Figure 11.

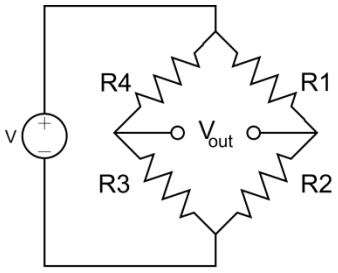


Figure 11: DC Wheatstone bridge

Resistors are denoted as R1 through R4. Usually at least one of the resistors, and in some cases as many as all four resistors are transducers whose resistance changes with applied stresses. The bridge is powered by a constant voltage controlled power source, V. The output

from the bridge is a DC signal that is proportional to the values of the resistors. The input output relationship is shown in Equation (4) [Wheeler 2004].

$$V_{out} = V_{in} \frac{R_3 R_1 - \mathrm{R}_4 R_2}{(R_2 + R_3)(R_1 + R_4)} \tag{4}$$

When the numerator equals zero, the bridge is considered balanced and the output is negligible. As applied stresses cause changes in the values of the resistors, very small changes in resistance will be measurable as an output from the bridge. After manipulating equation (4) the changes in resistance of each resistor can be related to the output from the bridge. Note the inclusion of a non-linear term for large changes in resistance. If changes in resistance are small the non-linear term can be neglected. In the general case each change in resistance is assumed to result from both mechanically applied stresses and increases in temperature. Changes in resistance caused by changes in temperature are related using equation (5).

$$\Delta \mathrm{R} = \alpha_{\mathrm{TCR}} \Delta \mathrm{T} \tag{5}$$

Where α_{TCR} is the thermal coefficient of resistance. Ceramic resistors typically used on circuit boards are designed to have a low thermal coefficient of resistance (100ppm/°C). Pure metals have much higher coefficients of thermal resistance typically around 4000 ppm/°C that must be accounted for when using a Wheatstone bridge to measure changes in resistance due to applied mechanical stresses.

3.1.1 Temperature Compensation

If $V_{out} = 0 \rightarrow R_1 R_3 = R_2 R_4$ which is known as the balance condition. When the resistors in the bridge change:

$R_1 \rightarrow R_1 + \Delta R_1$

$R_2 \rightarrow R_2 + \Delta R_2$

$R_3 \rightarrow R_3 + \Delta R_3$

$R_4 \rightarrow R_4 + \Delta R_4$

Plugging these values in for the above equation:

$$\Delta V_{out}|_{V_{out}=0} = \frac{V_{in} r}{(1+r^2)}\left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2} + \frac{\Delta R_3}{R_3} - \frac{\Delta R_4}{R_4}\right](1-\eta) \tag{6}$$

Where

$$\eta = \frac{1}{\left[1 + \left(\frac{1+r}{\frac{\Delta R_1}{R_1} + \frac{\Delta R_4}{R_4} + r\left(\frac{\Delta R_2}{R_2} + \frac{\Delta R_3}{R_3}\right)}\right)\right]} \tag{7}$$

If $r = {R_2}/{R_1} = 1 \; and \; \Delta R_1 < 0.02$,then the error from η is approximately 1%.

For a double arm Wheatstone bridge if $\Delta R_1 > 0 \; and \; \Delta R_3 = \Delta R_4 = 0$

$$\Delta V_{out} = \frac{V_{in} r}{(1+r^2)}\left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2}\right](1-\eta) \tag{8}$$

For the case where R1 is a tranducer mounted on a specimen, and R2 is a reference gage that does not feel any mechanical stress, then $\Delta R_2 = 0$ unless $\Delta T \neq 0$.

Changes in resistance can be related to strain imparted on a resitor by a gage factor. A gage factor is defined as:

$$S_g^{mech} \equiv \frac{{dR}/{R}}{\varepsilon_{mech}} \tag{9}$$

Similarily a gage factor can be calculated for strain caused by thermal expansion.

$$\Delta L = \alpha L_o \Delta T \tag{10}$$

$$\varepsilon_{thermal} = \frac{\Delta L}{L_o} = \alpha_{exp}\Delta T \tag{11}$$

$$S_g^{expand} \equiv \frac{dR/R}{\varepsilon_{thermal}} \tag{12}$$

Where α is the thermal coeficient of expansion. As discussed before resistance can also be effected by variations in ambient temperature. If R1 = R2 and are identical materials, then the change in resistance from the thermal coeficient of resistance is the same for both materials when subjected to the same ΔT.

$$\Delta R_1^{\Delta T} = \Delta R_2^{\Delta T} = \alpha_{TCR}\Delta T \tag{13}$$

Changes in resistnace are related to the mechanical and thermal strain felt by each resistor through a gage factor, Sg, that is generaly unknown (and possibly non-linear) unless the resistors are comercial strain gages. When solder joints are used as resistors the gage factor is probably non-linear in nature.

$$\frac{\Delta R_1}{R_1} = S_{g1}^{mech}\varepsilon_{mech} + S_{g1}^{expand}\varepsilon_{thermal} + \Delta R_1^{\Delta T} \tag{14}$$

$$\frac{\Delta R_2}{R_2} = S_{g2}^{expand}\varepsilon_{thermal} + \Delta R_2^{\Delta T} \tag{15}$$

When $\Delta T \neq 0$, the thermal terms inside the brackets cancle each other

$$\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2} = S_{g1}\varepsilon_{mech} + S_{g1}\varepsilon_{thermal} + \alpha_{TCR}\Delta T - S_{g2}\varepsilon_{thermal} - \alpha_{TCR}\Delta T = S_{g1}\varepsilon_{mech} \tag{16}$$

This setup has the effect of negating the effects of fluxuating temperatures if R1 is feeling a mechanical strain and R2 is not stressed mechanicaly, but undergoes a change in resistance due to changing temperatures. Both R1 and R2 must be relativley near each other so that they track changes in resistance from the ΔT simulaneously.

Similarily to the development above, if R3 and R4 are indentical conductors and also experiencing changes in resistance from fluctuations in the ambient temperature, but no external forces, then their contributions to the output of the bridge will cancle each other out in a manner similar to the double arm bridge.

$$V_{out} = \frac{V_{in} r}{(1+r^2)}\left[\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2} + \frac{\Delta R_3}{R_3} - \frac{\Delta R_4}{R_4}\right](1-\eta) \tag{17}$$

$$\frac{\Delta R_1}{R_1} = S_{g1}^{mech}\varepsilon_{mech} + S_{g1}^{expand}\varepsilon_{thermal} + \Delta R_1^{\Delta T} \tag{18}$$

$$\frac{\Delta R_2}{R_2} = S_{g2}^{expand}\varepsilon_{thermal} + \Delta R_2^{\Delta T} \tag{19}$$

$$\frac{\Delta R_3}{R_3} = S_{g3}^{expand}\varepsilon_{thermal} + \Delta R_3^{\Delta T} \tag{20}$$

$$\frac{\Delta R_4}{R_4} = S_{g4}^{expand}\varepsilon_{thermal} + \Delta R_4^{\Delta T} \tag{21}$$

$$\begin{aligned}\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2} + \frac{\Delta R_3}{R_3} - \frac{\Delta R_4}{R_4} \\ = S_{g1}^{mech}\varepsilon_{mech} + S_{g1}^{expand}\varepsilon_{thermal} + \Delta R_1^{\Delta T} - S_{g2}^{expand}\varepsilon_{thermal} \\ + \Delta R_2^{\Delta T} + S_{g3}^{expand}\varepsilon_{thermal} + \Delta R_3^{\Delta T} - S_{g4}^{expand}\varepsilon_{thermal} + \Delta R_4^{\Delta T}\end{aligned} \tag{22}$$

In the situation where very small changes in resistance need to be measured in specimens with very small absolute resistances the DC Wheatstone bridge may not be the best choice. For example Figure 12 shows the equivalent circuit for a DC bridge. The power that must be dissipated by the circuit is inversely proportional to the equivalent resistance of the circuit. If sufficiently large enough resistors are not used in the bridge the circuit will quickly overheat.

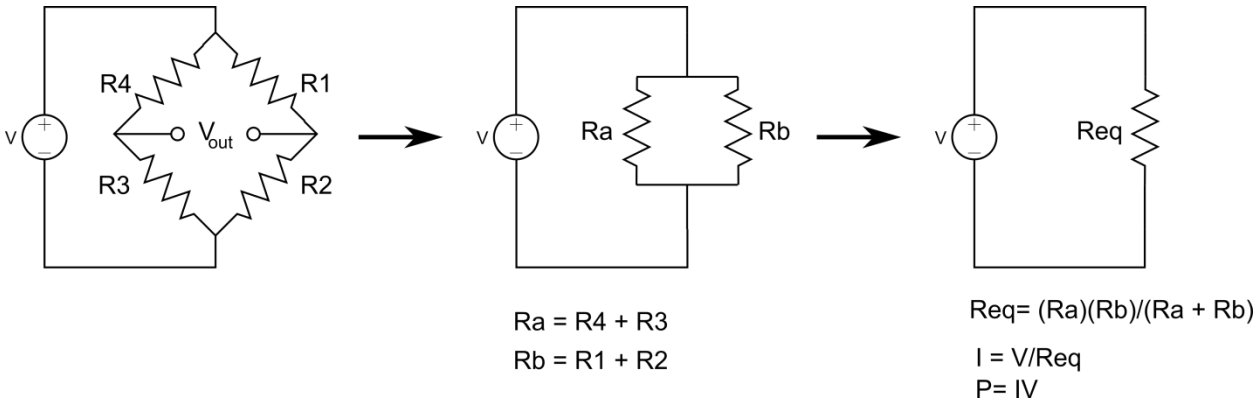


Figure 12: Equivalent circuit for a DC Wheatstone bridge

Because of the overheating problem a shunt resistor, Rs, may be used to increase the equivalent resistance of a DC bridge when measuring a small resistance, Rx, as in Figure 13. In the shunt resistor configuration R2 is equal to the sum of Rs and Rx.

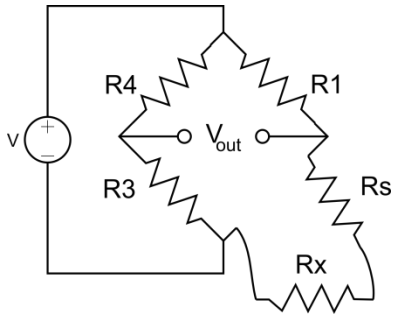


Figure 13: Addition of a shunt resistor to facilitate measurement of changes in resistance of the small resistance Rx

$$V_{out} = \frac{1}{4}\frac{\Delta R_2}{R_2}V_{in} \tag{23}$$

$$Let:\ R_2 = R_x + R_s \tag{24}$$

$$Let:\ \beta = \frac{R_s}{R_x} \tag{25}$$

$$V_{out} = \frac{1}{4}\frac{\Delta R_x}{(R_x + R_s)}V_{in} = \frac{1}{1+\beta}\left[\frac{1}{4}\frac{\Delta R_x}{R_x}V_{in}\right] \tag{26}$$

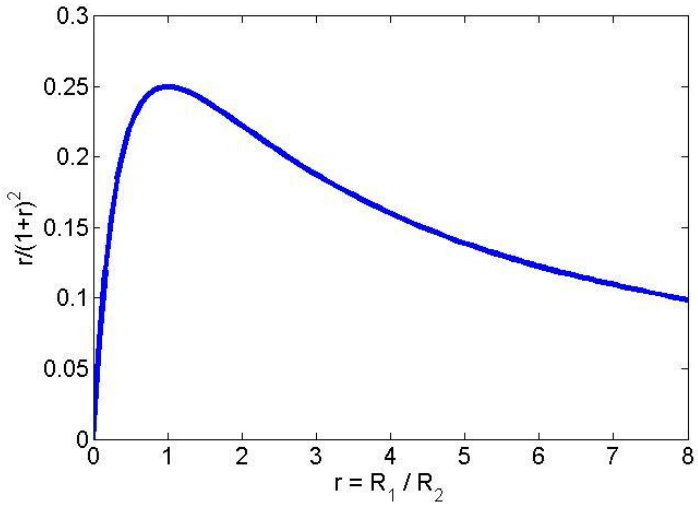


Figure 14: Sensitivity of the DC Wheatstone bridge for different choices of resistors

If the output from the bridge is only a function of R2, then the effect of the shunt resistor on the sensitivity of the bridge can be developed in equations (24) through (26) and illustrated in Figure 14. In response to the shortcomings of the DC bridge the resistance spectroscopy technique uses a modified Wheatstone bridge to make precise measurements on small changes in resistance on specimens with a small absolute value of resistance.

3.2 Alternating Current Wheatstone Bridge

An alternating current Wheatstone bridge behaves in a similar manner to its DC counterpart. The advantage of using an AC bridge is the ability to monitor both phase and magnitude information from the signal leaving the bridge which provides richer information about the device being monitored.

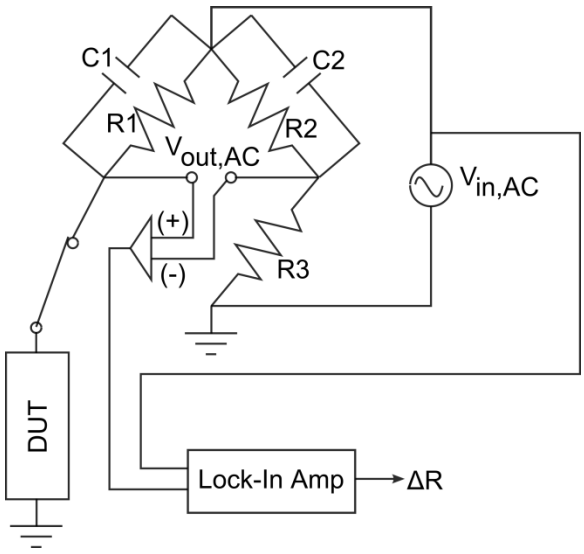


Figure 15: Schematic of AC Wheatstone Bridge

Figure 15 shows a schematic of an AC Wheatstone bridge. The device under test (DUT) is a resistive transducer. In addition to the four resistive elements found in the DC bridge, the AC bridge has two additional capacitors (C1 and C2) and is powered by a sinusoidal power source. Both the input into the bridge, and the output of the bridge are provided to a lock-in-amplifier which implements phase sensitive detection. Phase sensitive detection is the topic of the next section. Because the input to the AC bridge is sinusoidal, the output will be a sinusoidal signal of the same frequency as the input signal. If the change in resistance is small, the magnitude of the output signal will be small, and there may be a phase shift between the input and output signal. Capacitors C1 and C2 help eliminate any stray inductances that occur from wires running between the DUT and the test circuit.

The output equation for the AC Wheatstone bridge is very similar to the DC bridge, except impedances must be used to describe the components.

$$V_{out} = V_{in}\frac{Z_3Z_1 - Z_4Z_2}{(Z_2 + Z_3)(Z_1 + Z_4)} \tag{27}$$

Using the following definitions for impedances we can write our transfer function in terms of the discrete components in the bridge.

Table 3: List of impedances for passive electrical components used in an AC Wheatstone bridge

Device	Impedance/Resistance
Inductor	$Z_L = j\omega L$
Capacitor	$Z_C = -j\frac{1}{\omega C} = \frac{1}{j\omega C}$
Resistor	$Z_R = R$

Z_1 is the impedance of the C_1 and R_1 in parallel. R4 is the resistance of the DUT. Our transfer function is simply the ratio of the output to the input.

$$\frac{V_{out}}{V_{in}} = \frac{Z_3 Z_1 - Z_4 Z_2}{(Z_2 + Z_3)(Z_1 + Z_4)} \tag{28}$$

Starting with the Z_1 term.

$$Z_1 = \frac{1}{j\omega C_1 + \frac{1}{R_1}}$$

The Z_2 term is calculated the same as the Z_1 term but with the subscripts changed, and the impedance of a resistor is simply $Z_3 = R_3$ and $Z_4 = R_4$.

Converting to the frequency domain:

$$j\omega \to s$$

$$Z_1(s) = \frac{1}{sC_1 + \frac{1}{R_1}} = \frac{{}^{1}/_{C_1}}{s + \frac{1}{C_1 R_1}}$$

$$Z_2(s) = \frac{1/C_2}{s + \frac{1}{C_2 R_2}}$$

$$Z_3(s) = R_3$$

$$Z_4 = R_{pkg}$$

Plugging in and expanding the transfer function results in equation (29).

$$H(s) = \frac{Z_3 Z_1 - Z_4 Z_2}{(Z_2 + Z_3)(Z_1 + Z_4)} = \frac{\left(C_2 R_1 R_2 R_3 - C_1 R_1 R_2 R_{pkg}\right)s + R_1 R_3 - R_2 R_{pkg}}{\left(C_1 C_2 R_1 R_2 R_3 R_{pkg}\right)s^2 + \left(C_2 R_1 R_2 R_3 + C_1 R_1 R_2 R_{pkg} + C_1 R_1 R_3 R_{pkg} + C_2 R_2 R_3 R_{pkg}\right)s + \left(R_1 R_2 + R_1 R_3 + R_2 R_{pkg} + R_3 R_{pkg}\right)} \quad (29)$$

Table 4: Typical values used in an AC Wheatstone bridge

Component	Value
C1, C2	0.01uf = 1e-8 f
R1,R2	10 ohms
R3	3 ohms
DUT	3.1 ohms

Using the transfer function in equation (29) and the values listed in Table 4, the resulting bode plot is shown in Figure 16.

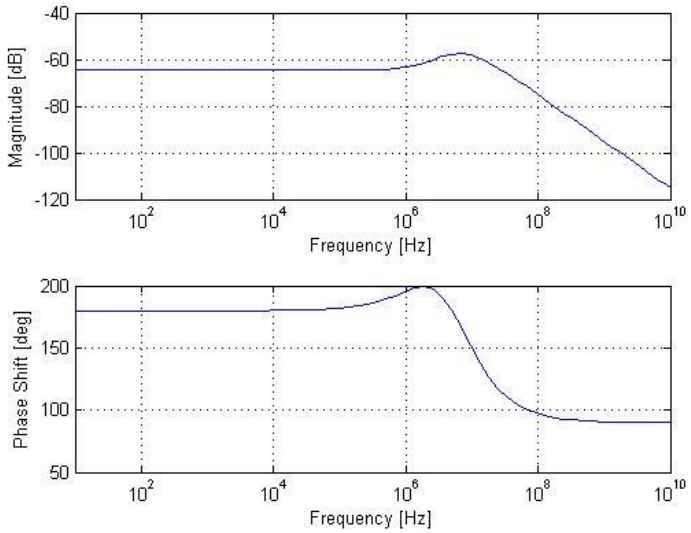


Figure 16: Bode plot for AC Wheatstone bridge

The bode plot demonstrates that the output signal for a small change in resistance will be very small, which motivates the use of phase sensitive detection described in the next section.

3.3 Phase Sensitive Detection

Phase sensitive detection (PSD) is an elegant technique that is well suited for measuring the magnitude and phase of small signal corrupted with noise. As described in the previous section, the output from an AC Wheatstone bridge is expected to be a sinusoidal signal with the same frequency as the input signal but much smaller in magnitude and possibly phase shifted. The driving signal input into the top of the AC bridge is denoted as the reference signal. The reference signal, $V_{in,AC}$, is a very clean sine wave. The output from the AC bridge, $V_{out,AC}$, is a very small magnitude signal corrupted with noise greater than the magnitude of the signal. For the sake of describing the process of performing PSD assume that both the reference signal, V_{REF}, and the measured signal, V_{SIG}, are pure sine waves.

$$V_{REF} = V_1 \sin(\omega_1 t + \theta_1) \tag{30}$$

$$\mathrm{V_{SIG}} = V_2 \sin(\omega_2 t + \theta_2) \tag{31}$$

$$\mathrm{V_M} = (V_{REF})(V_{SIG}) \tag{32}$$

$$\mathrm{V_M} = \frac{V_1 V_2}{2}[cos([\omega_1 - \omega_2]t + \theta_1 - \theta_2) - cos([\omega_1 + \omega_2]t + \theta_1 + \theta_2)] \tag{33}$$

$$\mathrm{V_M} = \frac{V_1 V_2}{2}[cos(\theta_1 - \theta_2) - cos(2\omega_1 t + \theta_1 + \theta_2)] \tag{34}$$

$$\mathrm{V_{PSD}} = \frac{V_1 V_2}{2} cos(\theta_1 - \theta_2) \tag{35}$$

In equation (32) the signals are multiplied together. In equation (33) the product to sum trigonometry identify is applied to $\mathrm{V_M}$. The frequency of the reference and measured signal are identical, so equation (33) reduces to equation (34). If a low pass filter with a cutoff less than $2\omega_1$is applied to equation (34) the higher frequency terms are attenuated to a negligible level as shown in equation (35). Equation (35) is the phase sensitive detection signal and is a constant value proportional to the magnitudes of the reference and measured signal and the phase shift between the signals. The values of V1 and θ_1 are known quantities. The unknown values are V2 and θ_2. To create a second independent equation that allows both unknown values to be solved for, the reference signal is shifted by 90 degrees, i.e. $\mathrm{V_{REF,2}} = V_1 \sin(\omega_1 t + \theta_1 + 90°)$. Repeating the process in equations (32) through (35) a similar relation can be determined for the shifted reference signal.

$$\mathrm{V_{PSD,2}} = \frac{V_1 V_2}{2} sin(\theta_1 - \theta_2) \tag{36}$$

With two independent equations all the unknowns related to the measured signal can be solved from equations (35) and (36). For historical reasons the results are typically communicated as a phasor [Valkenburg 1995].

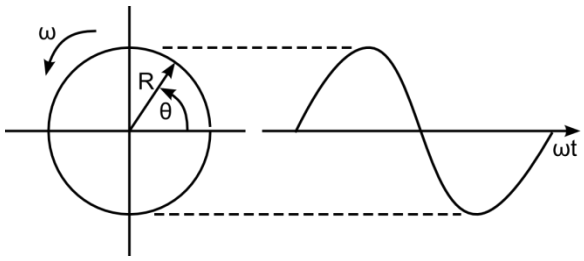


Figure 17: Phasor representation of a sine wave

In the more realistic case the measured signal is corrupted with high frequency noise as shown in equation (37). Fortunately the high frequency terms will all be greater than the $2\omega_1$ term in equation (34) and will also be attenuated to negligible levels after the lowpass filter.

$$V_{SIG} = V_{2,1}\sin(\omega_{2,1}t + \theta_{2,1}) + V_{2,2}\sin(\omega_{2,2}t + \theta_{2,2}) + V_{2,3}\sin(\omega_{2,3}t + \theta_{2,3}) + \cdots \quad (37)$$

$$\omega_{2,1} \ll \omega_{2,2};\ \omega_{2,2} < \omega_{2,3}; \ldots \quad (38)$$

Functionally phase sensitive detection (shown in Figure 18) is easy to implement. The reference and measured signals from the AC Wheatstone bridge are multiplied together and then passed through a low pass filter. In practice the multiplication is handled using a digital signal processing (DSP) integrated circuit. The low pass filter is typically a higher order filter such as an 8th order Butterworth filter. In Figure 18 Vm is the multiplied signal, R_f and C_f are components in the low pass filter, and V_{psd} is the phase sensitive detection result.

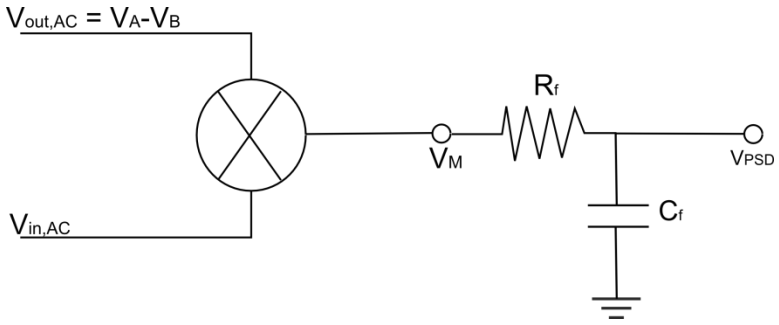


Figure 18: Functional Schematic of Lock in Amplifier

The phase sensitive detection method is implemented in a commercially available lock in amplifier (Figure 19). For this work an SRS830 lock-in amplifier from Stanford Research Systems (1290-D Reamwood Avenue, Sunnyvale, CA 94089) was used to perform phase sensitive detection.

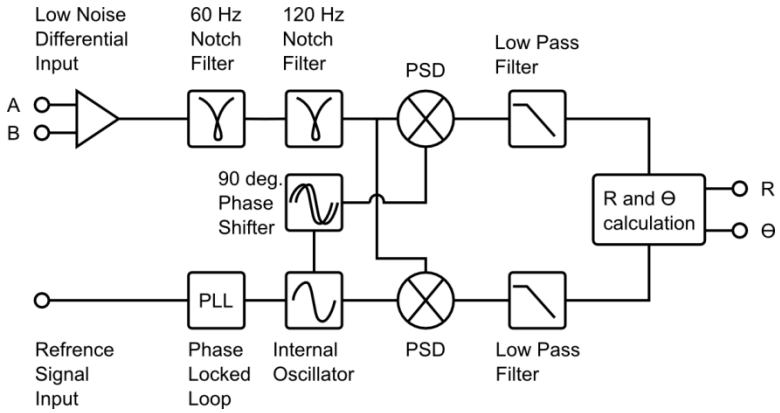


Figure 19: Simplified block diagram adapted from SRS830 Lock-in amplifier user manual

3.4 Theory vs. Observation at Different Frequencies

For historical reasons the data presented in this document used the same AC Wheatstone bridge for a number of different data sets. When an industrial partner became interested in implementing the RS technique in a production grade product there was a need implement the technique at a lower frequency to minimize the cost of the measurement hardware.

Table 5: Component values for a low frequency AC bridge

Component	Value
C1, C2	10 uF
R1,R2	10 ohms
R3	11.3 ohms
DUT	11.35 ohms

Based on the transfer function we expect the circuit to behave according to the bode plot shown in Figure 20.

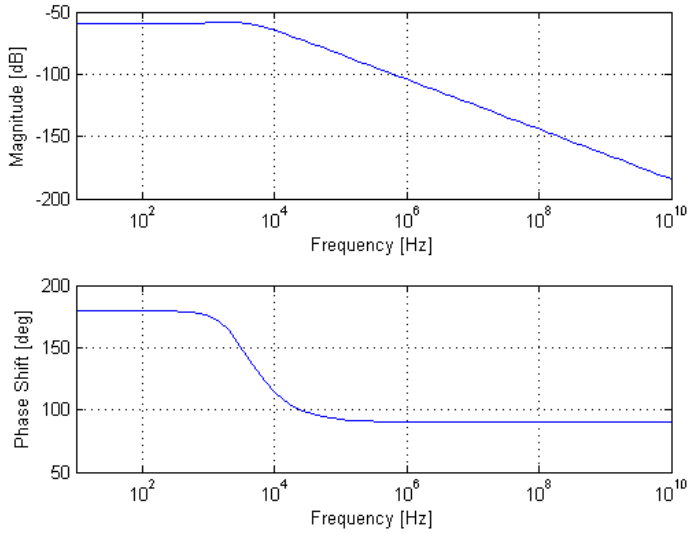


Figure 20: Bode plot from a circuit optimized for a low frequency measurement

A labview program was created to automatically iterate between three different driving frequencies for the AC Wheatstone bridge during a test and record the resulting outputs. The device under test was the daisy chain resistance of a PBGA324 component. The test environment was a random vibration profile. Full details about the test setup can be found in the resistance spectroscopy section. As the package became damaged the resistance of the device under test increases. Data was recorded for a driving frequency of 100, 1000, and 10000 Hz. The raw magnitude data from these measurements is presented in Figure 21. A zoomed view of the same data set around the time of interest is shown in Figure 22. The data is reported as the R component, or magnitude, of the AC signal measured by the lock-in amplifier.

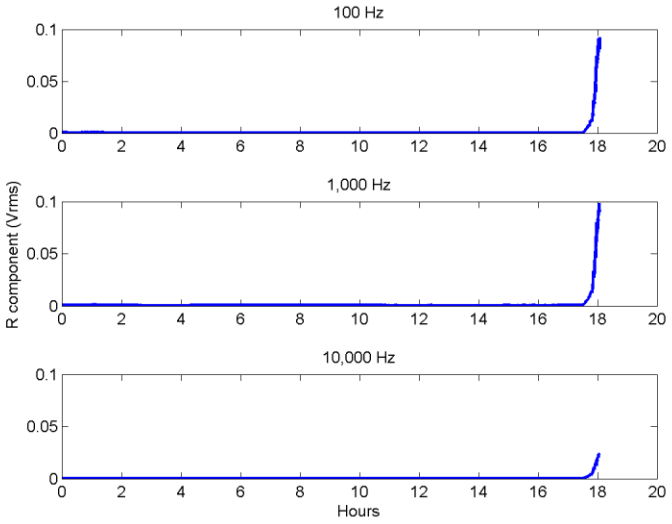


Figure 21: Magnitude component recorded at different driving frequencies for the same test

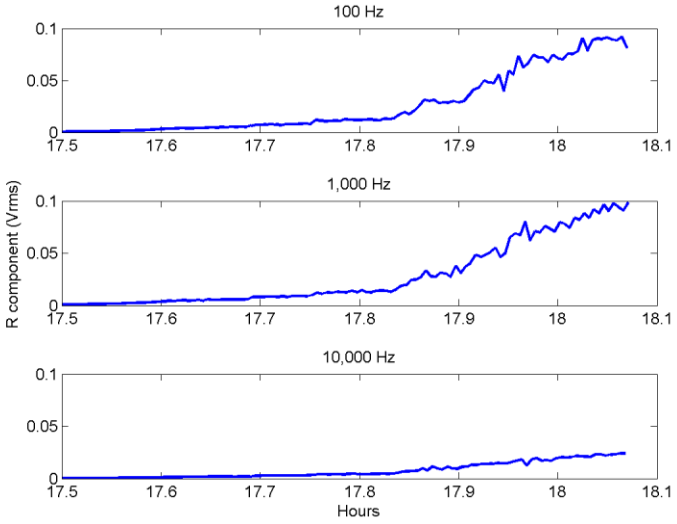


Figure 22: Zoomed view aruond the region of interest from Figure 21

As predicted by the bode plot the sensitivity of the measurement at 1 kHz is just slightly better than at 100 Hz. The drop off in sensitivity predicted by the bode plot can be seen for the data recorded at 10 kHz. These results confirm that the RS technique will detect small changes in resistance at different driving frequencies as predicted by the theoretical transfer function.

4 Risk Based Decision Making

Risk based decision making is a methodology that embodies how to make prognostic predictions of failure useful for the end user of a system. Using a process established in [Engle 2009], statistically defendable methods for using predictions and taking action based on those predictions is established. Using the same remaining useful life prediction in different ways, risk based decision making can be structured to optimize, safety, cost, availability, and logistics.

4.1 Methodology

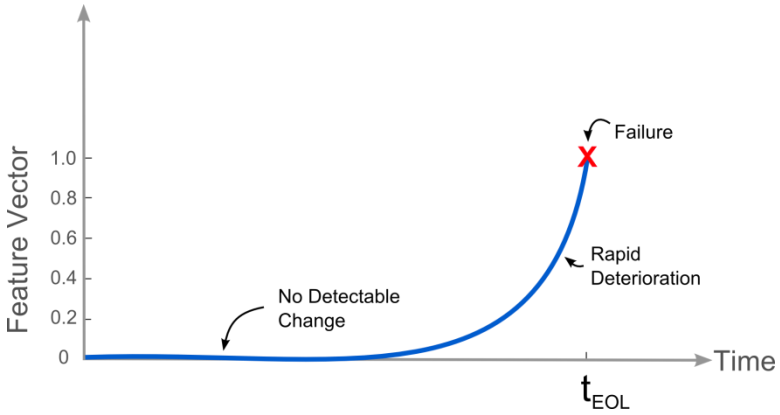


Figure 23: Typical wear out profile of a safety and mission critical system

Figure 23 shows the hypothetical wear out profile of a system. The y-axis represents the value of a feature vector, i.e. some characteristic of the system being monitored to characterize wear. Risk based decision making and the feature vector is generic and can include any salient feature of the system, but this work will mainly focus on feature vectors derived from the resistance spectroscopy measurements. A typical component will show no significant change for a large percentage of its life. Towards the end of a components life the system tends to rapidly deteriorate until failure occurs. The end of life (EOL) condition is easy to identify, the system no longer works as expected. Failure is much more difficult to quantify because “operating as

expected" can depend on many different interpretations. For the purposes of developing prognostic algorithms, a binary representation is preferred as stated in equation (36).

$$C_{EOL}(x(t)) = \begin{cases} 1, if\ EOL\ is\ reached \\ 0, otherwise \end{cases} \quad (39)$$

The link between a monitored feature of the system wearing out, and the failure of the system requires a correlation between the two variables. This is most easily represented as a failure threshold (Figure 24).

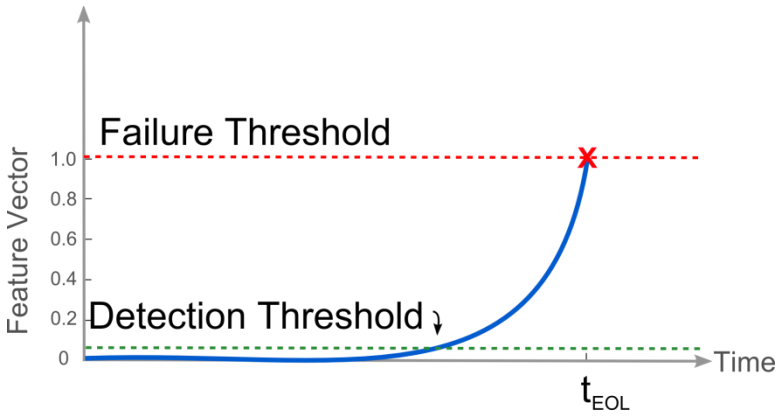


Figure 24: Failure threshold and detection threshold that facilitates quantifying failures based on the value of a feature vector

When monitoring a feature vector in the regime where nothing is changing it is not prudent to try to predict the end of life for the component. Therefore a detection threshold can be established which matches with the noise floor of your measurement. Until the feature vector increases above the detection threshold the system is considered healthy and no predictions of end of life are made. After the feature vector has broken through the detection threshold predictions are made to predict the time in the future when the feature vector will cross the failure threshold and the system will suffer a failure (C_{EOL} = 1).

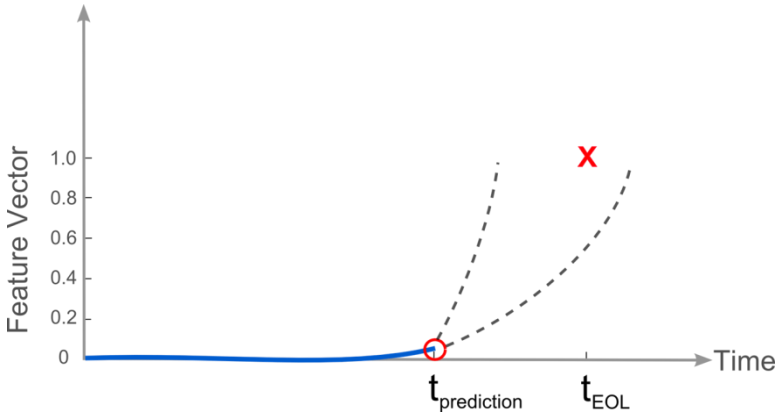


Figure 25: The prediction challenge, estimating the time until the feature vector crosses the failure threshold

Figure 25 illustrates the challenge of prognosticating failure. Given a limited amount of information about a feature vector, the challenge is to predict when in time the feature vector will cross the detection threshold which correlates with failure of the system. Many hard to quantify variables such as future usage, stress profiles, interacting stress conditions, and component specific material defects will affect the manner in which the feature vector progresses to the failure threshold. With so many uncertainties a range of possible predicted failure times can be supplied to the end user (Figure 26). Remaining useful life (RUL) is denoted as the time between the prediction and the predicted failure.

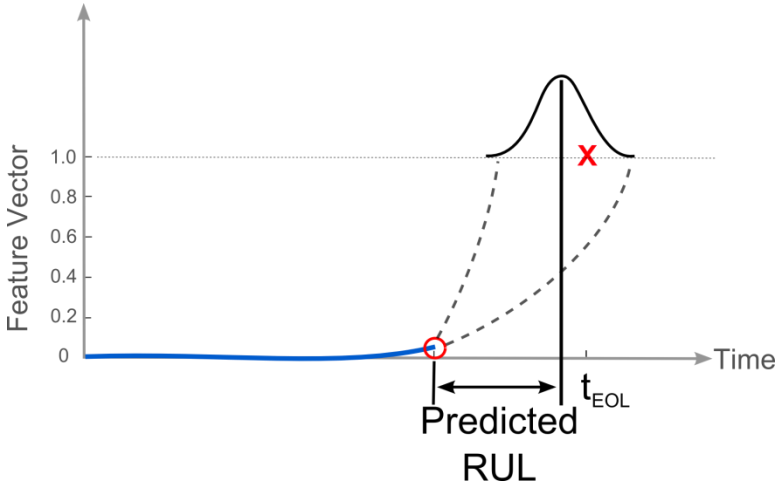


Figure 26: Uncertainty in the failure prediction

A number of different methods for propagating the current state of a feature vector forward in time until the failure threshold is crossed will be discussed in this document. The predicted failure distribution is drawn as a Gaussian distribution for the sake of illustration, but is not restricted to being described by closed form statistics. Unfortunately making predictions is not the same as providing useful information to an end user. To practice risk based decision making using inputs from an algorithm that monitors a feature vector and predicts, or prognosticates, failure you need to establish some high level guidelines for use of the system.

Risk based decision making requires that the user of a system define acceptable levels of risk they are willing to tolerate. The user must define the following inputs.

4.1.1.1 Maximum Allowable Probability of Failure

This quantity represents the user's tolerance for an unplanned failure. In a safety or mission critical application this probability would be very small.

4.1.1.2 Maximum probability of proactive maintenance

This probability represents the tolerance for replacing a system before it has failed. This value has a direct impact on the cost of operating a mission and safety critical system. By definition for a critical system this value should be a very low probability, but the practicality of limited resources dictates that this probability be as large as possible to minimize operating cost.

4.1.1.3 Required Lead Time for Taking Action

The required lead time for taking action encompasses the logistical realities associated with operating and maintaining complicated systems. Due to finite resources, it may not be possible to keep in stock every spare part for a system. Other constraints may be the availability of properly trained personnel to perform maintenance actions.

4.1.1.4 Required Confidence in Predictions

It is important to note that predictions are uncertain due to the wide number of variables that affect the rate of wear imparted on a system. The required confidence in a prediction can be used to avoid taking action when information from the prognostic algorithms may not be fully converged to the correct failure time.

The inputs for making risk based decision making are not necessarily hard constraints. The more physical justification that can be applied to the inputs the better the resulting decision making performance will be, but again many of these inputs will have to be simply estimated. Using a statistically defendable approach based on the true state of the system is of course preferable to the standard procedure of using traditional reliability methods to schedule replacement of worn out components.

Using the inputs provided by the user for risk based decision making a window of opportunity can be identified for taking action based on the predicted EOL. The area under a probability density function is defined as unity. Therefore a percentage of the area under the PDF represents a probability of an event occurring. The maximum allowable probability of proactive maintenance can be subtracted from the maximum probability of failure to determine a window of opportunity where it is appropriate to take action (Figure 27). A hypothetical example is using risk based decision making to schedule the appropriate time to perform maintenance (Figure 28).

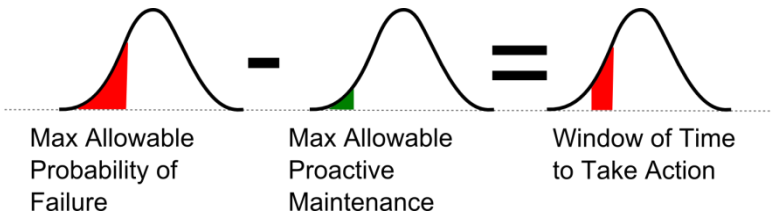


Figure 27: Window of opportunity for taking action

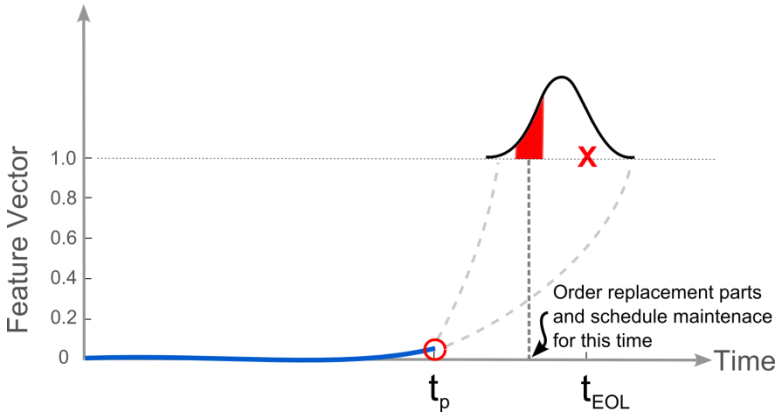


Figure 28: Appropriate time to schedule maintenance

If new information becomes available after an initial prediction of RUL (or EOL) but before the actual failure, then it is advantageous for the prognostic algorithm to incorporate the latest available information and issue an updated prediction of failure. The process of incorporating new measurements of the feature vector and issuing revised predictions repeats until failure occurs.

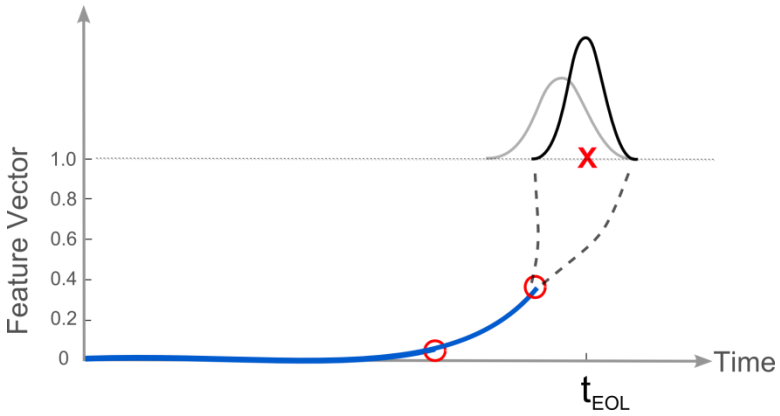


Figure 29: Updated failure prediction after new information is available

4.2 Conclusion

A framework for risk based decision making has been presented in this section. Statistically defendable methods for using information output from a prognostic algorithm to take action in the presence of uncertainty were described. The presented framework, known as risk based decision making, was introduced in general terms, but will be implemented for the

case of monitoring the health of electronics in mission and safety critical electronics. The recursive nature of updating predictions as new information becomes available motivates the use of recursive filters described in the next section.

5 Demonstration of Recursive Filtering

Recursive filtering is a technique of combining a model of a systems expected dynamic behavior with noisy measurements to estimate the true state of a system. It is acknowledged that neither the system model nor the measurement are perfect, but rather have some amount of error associated with them. Due to the recursive nature of the algorithms they are suited to real time applications where a stream of data is constantly updated, such as tracking a leading indicator of failure. A number of different filtering algorithms have been implemented for this document, but they have been applied to different data sets and do not provide an easy comparison between the different methods. This section will be used to introduce the different filtering methods and will benchmark their performance against a standard test case. Because the test case consists of simulated measurements the ground truth is available. In practical situations presented later in this document the ground truth is not known and the algorithms must be evaluated based solely on their prognostic performance, which is not a fair comparison when discussing tracking performance. Due to the robust nature of the algorithms, poorly configured implementations will not necessarily appear different than a superior implementation unless a known good baseline is available. In this document the classic case of tracking a ballistic object upon re-entry to the earth's atmosphere will be discussed. The derivation and justification for the algorithms is included.

5.1 Least Squares Approaches

It is convenient to think of the recursive filtering algorithms as related to a least squares curve fitting problem. Many of the derivations of the recursive algorithms described in this

section are beyond the scope of the document, but a comparison between batch and recursive least squares algorithms provides intuition into the problem and is easily demonstrated.

5.1.1 Batch Least Squares

Fitting data to a model in the least squares sense traditionally implies the use of a batch least squares algorithm. In a batch least squares every data point must be available before processing can begin. Following the example in [Zarchan 2000], hypothetically say you were trying to estimate the true value of a noisy process. In this simple example you are making a noisy measurement of a constant value. The first assumption is to formally decide that you can model your system as a zero order polynomial.

$$\hat{\mathrm{x}}_{\mathrm{k}} = a_o \tag{40}$$

In the least squares method you desire to estimate a_o by minimizing the residual, R, between your estimate and actual observations, x_k^*.

$$R = \sum_{k=1}^{n} (\hat{\mathrm{x}}_k - x_k^*)^2 = \sum_{k=1}^{n} (a_o - x_k^*)^2 \tag{41}$$

Using calculus the optimum estimate of a_o can be obtained.

$$\frac{\partial R}{\partial a_o} = 0 = 2(a_o - x_1^*) + 2(a_o - x_2^*) + \cdots + 2(a_o - x_n^*) \tag{42}$$

$$0 = 2\left(na_o - \sum_{k=1}^{n} x_k^*\right) \tag{43}$$

$$a_o = \frac{\sum_{k=1}^{n} x_k^*}{n} \tag{44}$$

Equation (44) is the batch least squares solution to estimating the coefficients of a zero order polynomial. If the example was higher order than a zero order polynomial the answer

would not degenerate into the equation for the sample mean. This result will be used in the next section.

5.1.2 Recursive Least Squares

To formulate equation (44) as a recursive formula that does not need the entire data set to provide estimates of the coefficient of a zero order polynomial equation (44) is re-written.

$$\hat{\mathrm{x}}_{\mathrm{k}} = a_o = \frac{\sum_{i=1}^{k} x_i^*}{k} \tag{45}$$

Changing the subscripts to let k = k+1,

$$\hat{\mathrm{x}}_{\mathrm{k+1}} = \frac{\sum_{i=1}^{k+1} x_i^*}{k+1} = \frac{x_{k+1}^* + \sum_{i=1}^{k} x_i^*}{k+1} \tag{46}$$

By re-arranging equation (45),

$$k\hat{\mathrm{x}}_{\mathrm{k}} = \sum_{i=1}^{k} x_i^* \tag{47}$$

and then substituting (47) into (46)

$$\hat{\mathrm{x}}_{\mathrm{k+1}} = \frac{x_{k+1}^* + k\hat{\mathrm{x}}_{\mathrm{k}}}{k+1} \tag{48}$$

Adding $\hat{\mathrm{x}}_{\mathrm{k}} - \hat{\mathrm{x}}_{\mathrm{k}} = 0$ to the numerator of the previous equation,

$$\hat{\mathrm{x}}_{\mathrm{k+1}} = \frac{(k+1)\hat{\mathrm{x}}_{\mathrm{k}} + x_{k+1}^* - \hat{\mathrm{x}}_{\mathrm{k}}}{k+1} = \hat{\mathrm{x}}_{\mathrm{k}} + \frac{1}{k+1}(x_{k+1}^* - \hat{\mathrm{x}}_{\mathrm{k}}) \tag{49}$$

Changing the subscripts to k+1=k,

$$\hat{\mathrm{x}}_{\mathrm{k}} = \hat{\mathrm{x}}_{\mathrm{k-1}} + \frac{1}{k}(x_k^* - \hat{\mathrm{x}}_{\mathrm{k-1}}) \tag{50}$$

Equation (50) is the recursive least squares solution for estimating a zero order polynomial. For the trivial case of a zero order polynomial equation (50) is the formula for a

recursive average. If the $1/k$ term were replaced with a constant, $1/L$ term, the equation would become a moving average filter of length L.

A simulation shown in Figure 30represents the case of making a measurement of a signal corrupted by white noise. Figure 31 shows how the recursive least squares estimate (equation (50)) converges to the batch least squares method equation (44). The final estimate of the recursive method is exactly equal to the batch method. In situations where it is not possible to wait until all measurements are available, as in prognostics, recursive methods can be a useful tool to quickly and efficiently estimate properties of a system given a noisy measurement.

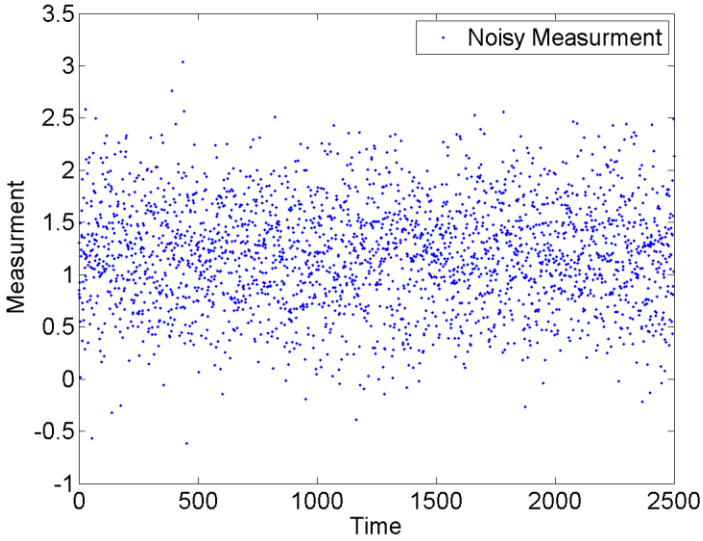


Figure 30: Simulation of a constant signal at a value of 1.2 measured in the presence of white noise

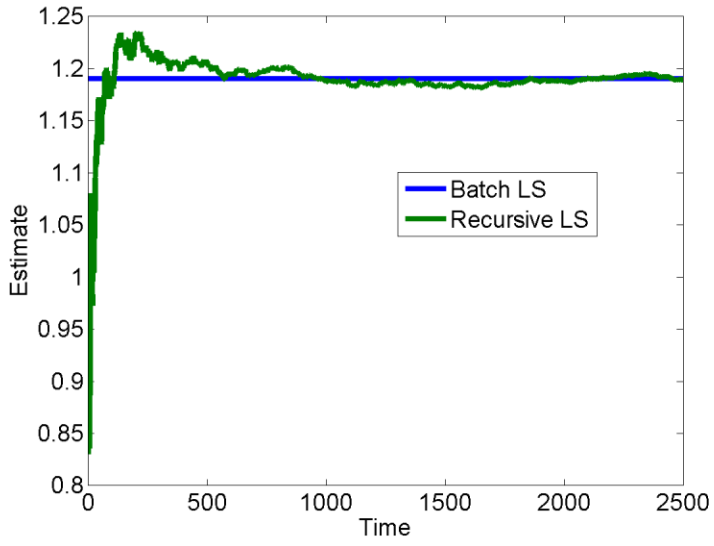


Figure 31: Comparison of the batch least squares approach to the recursive least squares approach

As the order of the polynomial that is assumed to describe the underlying system is increased the resulting least squares solutions become much more verbose. For example the equations for a first order system are given as equations (51)-(54) for the batch processing method, and (55)-(61) for the recursive algorithm.

$$\mathrm{y} = \mathrm{a}_0 + a_1 x \tag{51}$$

$$\mathrm{a}_0 = \frac{\bar{\mathrm{y}}\,\overline{x^2} - \bar{x}\,\overline{xy}}{n\overline{x^2} - (\bar{x})^2} \tag{52}$$

$$\mathrm{a}_1 = \frac{n\overline{xy} - \bar{x}\bar{y}}{n\overline{x^2} - (\bar{x})^2} \tag{53}$$

$$\bar{x} = \sum_{i=1}^{n} x_i \,;\, \bar{y} = \sum_{i=1}^{n} y_i \,;\, \overline{xy} = \sum_{i=1}^{n} x_i y_i \,;\, \overline{x^2} = \sum_{i=1}^{n} x_i^2 \tag{54}$$

$$\hat{\mathrm{x}}_\mathrm{k} = \mathrm{a}_0 + a_1 x \tag{55}$$

$$K_{1,k} = \frac{2(2k-1)}{k(k+1)} \tag{56}$$

$$\mathrm{K}_{2,k} = \frac{6}{k(k+1)T_s} \tag{57}$$

$$\bar{x}_k = \hat{x}_{k-1}^1 + \hat{x}_{k-1}^2 T_s \tag{58}$$

$$res_k = x_k^* - \bar{x}_k \tag{59}$$

$$\hat{x}_k^1 = \bar{x}_k + res_k K_{1,k} \tag{60}$$

$$\hat{x}_k^2 = \hat{x}_{k-1}^2 + res_k K_{2,k} \tag{61}$$

Equation (56) and (57) are denoted as gains, and they multiply the residual, res_k, between the state projection, $\bar{x}_k$, and the noisy measurement, x_k^* when calculating estimates of the polynomial coefficients.

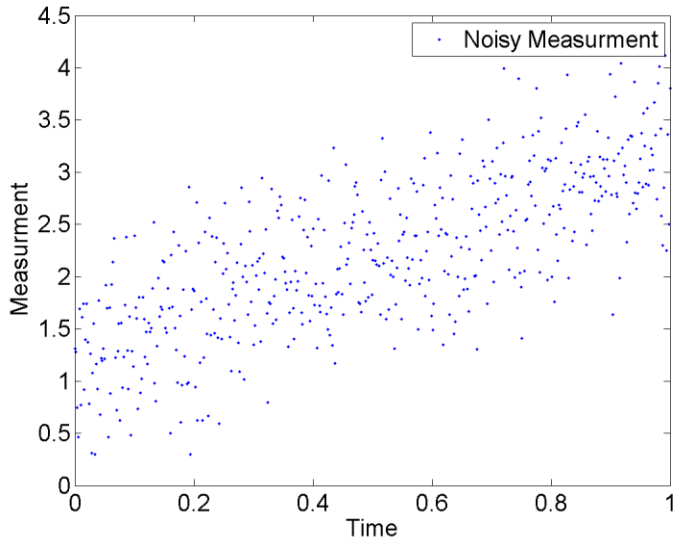


Figure 32: Simulated noisy data with an intercept of 1.2 and a slope of 2

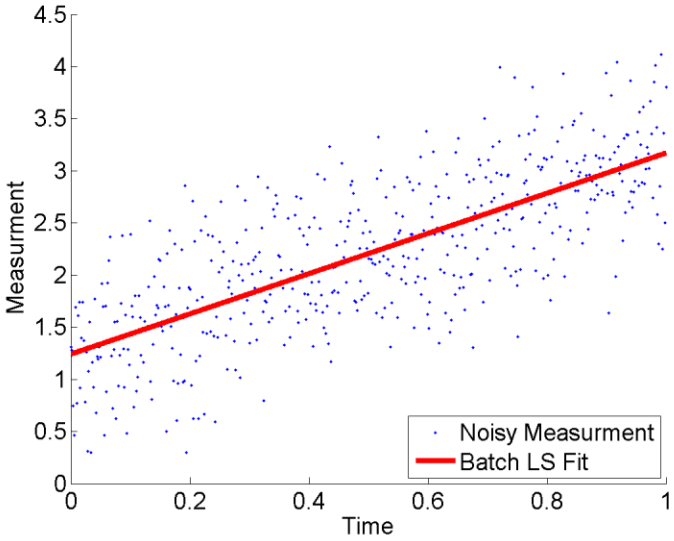


Figure 33: Batch least squares fit of noisy data

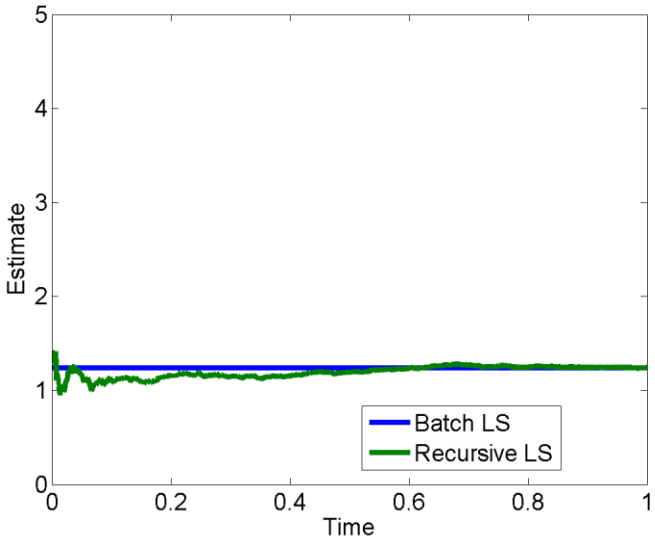


Figure 34: Comparison of batch least squares to recursive least squares for the a0 polynomial term

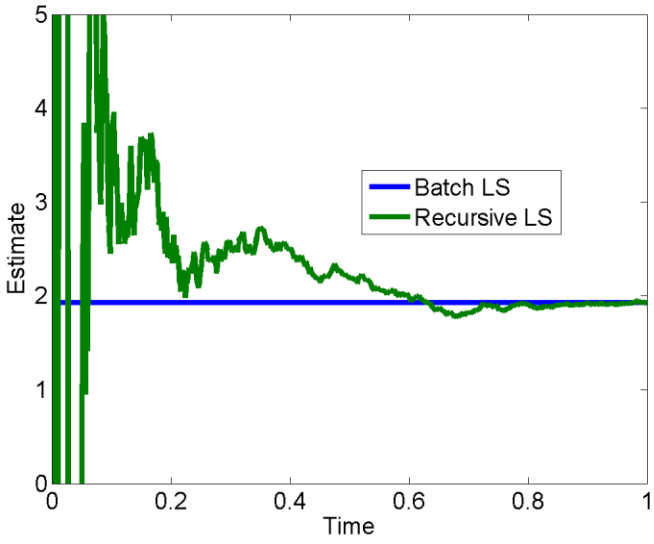


Figure 35: Comparison of batch least squares to recursive least squares for the a1 polynomial term

5.1.3 Savitzky-Golay Smoothing

Savitzky-Golay (SG) smoothing is a technique that is commonly applied to render visible the relative widths and heights of spectral lines in noisy spectrometric data [Vetterling 1992]. The technique is commonly known as Savitzky-Golay smoothing [Savitzky 1964, Steinier 1972 , Bromba 1981, Gorry 1990], least squares filtering, or digital smoothing polynomial filtering. The concept is to apply a window to a time series. The data point at the center of the window is smoothed by fitting a d-order polynomial to the data points inside the window of length N using a least squares approach. The best fit polynomial of order d, is evaluated at the center of the window to obtain a smoothed value for that point. The window is then moved forward one time step and the process is repeated. Following the formulation developed by [Orfanidis 1996] , to smooth a time series of noisy measurements x.

$$\mathrm{x} = [x_{-M}, \dots, x_{-1}, x_0, x_1, \dots, x_M] \tag{62}$$

For a window of size N, N=2M+1. Restrictions on N include that N be odd and that $\mathrm{N} \geq \mathrm{d} + 1$. A d order polynomial will be fit in a least squares sense to the measurements x.

$$\hat{\mathrm{x}}_{\mathrm{m}} = [c_o + c_1 m^1 + \cdots + c_d m^d] \tag{63}$$

Define d+1 polynomial basis vectors

$$\mathrm{s}_{\mathrm{i}}(m) = m^i \qquad -M \le m \le M \tag{64}$$

The N by (d+1) matrix S is comprised of columns s_i.

$$\mathrm{S} = [S_o, S_1, \ldots, S_d] \tag{65}$$

Smoothed values, $\hat{x}$, are evaluated as

$$\hat{\mathrm{x}} = \sum_{i=0}^{d} c_i s_i = [S_o, S_1, \ldots, S_d] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_d \end{bmatrix} = Sc \tag{66}$$

Then the following matrices are calculated;

$$\mathrm{F} = \mathrm{S}^{\mathrm{T}} S \tag{67}$$

$$G = SF^{-1} \equiv [g_0, g_1, \ldots, g_d] \tag{68}$$

$$B = SG^T \equiv [b_o, b_1, \ldots, b_d] \tag{69}$$

For a noisy sequence $x(n)$, the SG smoothed sequence is calculated as

$$y(n) = \sum_{m=-M}^{M} b_o(m) x(n+m) \tag{70}$$

Derivatives of the noisy sequence can be calculated using the G matrix.

$$y^{(i)}(n) = i! \sum_{m=-M}^{M} g_i(m) x(n+m) \tag{71}$$

A typical application of SG smoothing is shown in Figure 36. In many spectroscopy applications such as Auger or mass spectroscopy the location and width of a peak is corrupted with white noise. The challenge is to extract a signal representative of the underlying process with as much of the noise removed as possible. Moving average filters tend to fail on this type of

analysis since they attenuate high frequency components and create time shifts (Figure 37). In spectroscopy applications both the position and height of the peaks are important.

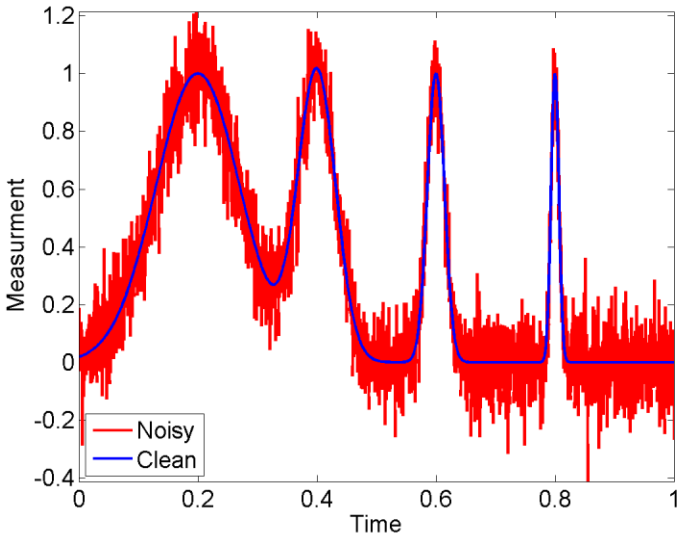


Figure 36: Synthesized spectrometric data set denoted as (Clean) and a noise corrupted version of the same signal (Noisy)

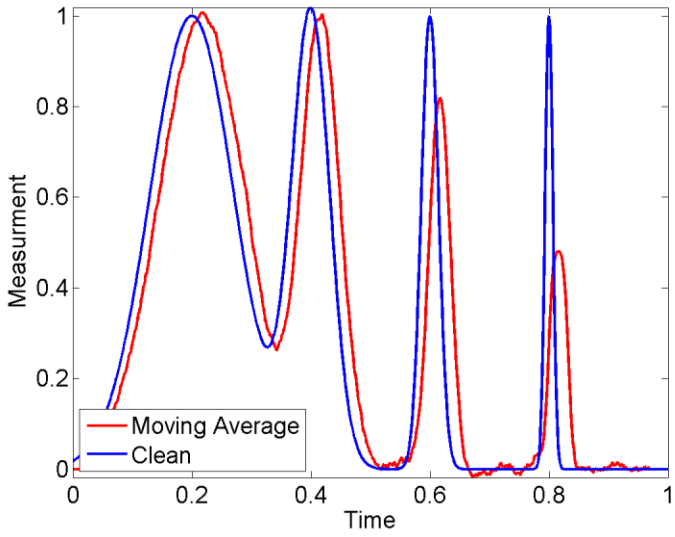


Figure 37: Moving average filter of length 65 applied to the synthetic data set

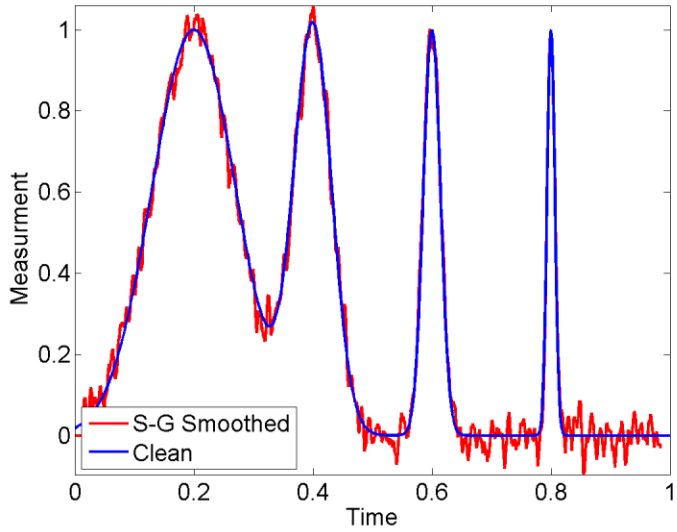


Figure 38: Savitzky-Golay smoothing applied to the noisy data set with a window length of 65 and a model order of 10

The SG method can also be used to estimate derivatives of noisy data. To illustrate why this is important a sinusoidal signal corrupted with white noise is synthesized. By calculating a numeric derivative using the rise over run method as shown in equation (72), the derivatives quickly become meaningless white noise. Notice the values of the vertical axis in Figure 39.

$$\left.\frac{dx}{dt}\right|_{numeric} = \frac{x_i - x_{i-1}}{t_i - t_{i-1}} \tag{72}$$

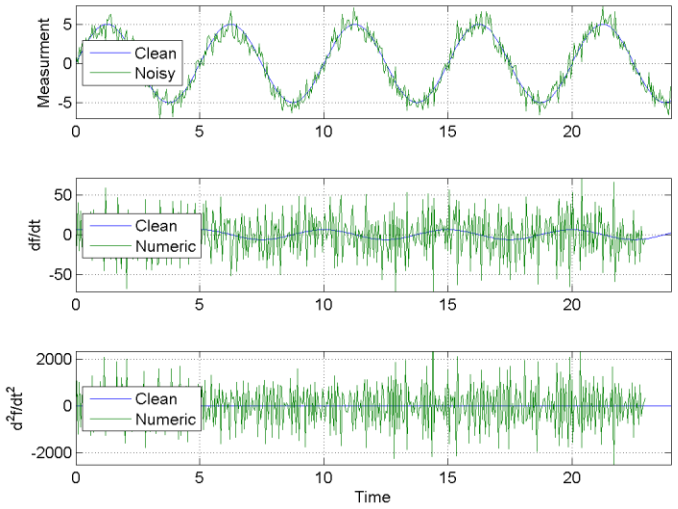


Figure 39: Numeric derivatives estimated for a sinusoidal signal corrupted with white noise

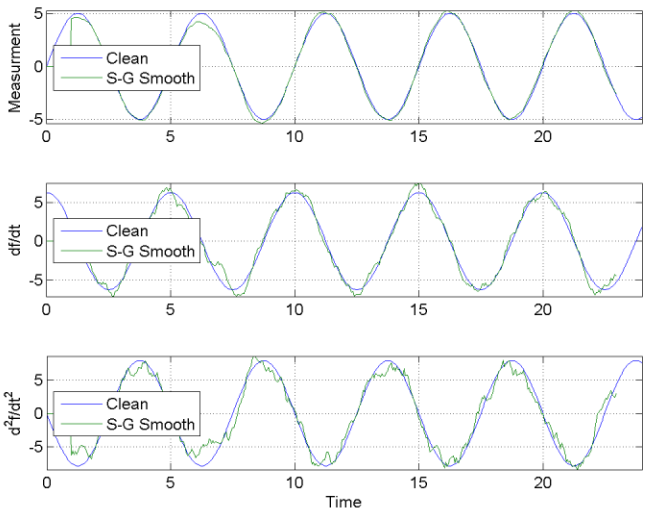


Figure 40: Savitzky-Golay estimates of derivatives of noisy data with a window of 41 and a model order of 3

In contrast to the numeric derivative example, the SG method provides smooth derivatives of higher derivatives. The first d derivatives are available with the SG method. Despite the premise of performing a least squares fit to every data window it is surprising that the SG filter can be implemented as efficiently as a FIR filter due to the fact that the filter coefficients can be calculated offline and could therefore be considered a real time method. Commonly in prognostics there is a need to smooth noisy measurements and to estimate smooth derivatives. Smooth derivatives are important because they are used in the prediction process and ultimately create smooth predictions of failure times. The three drawback of the SG method are the lag of length M between the most recently available data point, and the data point that is being smoothed. The window size and model order must be known *a priori* and the inability to quantify the uncertainty associated in the filter estimates of the signal and the derivatives.

5.2 State Estimator Approaches

State estimator approaches are a class of algorithms that are closely related to least squares problems, but that are formulated in a manner that is conducive to use in prognostic algorithms. The famous Kalman filter, and its related implementation the extended Kalman filter and particle filter are classified as state estimators. The simplest version of a Kalman filter will be derived in this section and the equations for implementing the other versions of the state estimator algorithms will be presented.

5.2.1 Kalman Filter

The Kalman filter was first described in [Kalman 1960]. The elegant notation from [Zarchan 2000] will be used to present the Kalman filter equations. The first assumption of the Kalman filter is the ability to model the system you are tracking as a set of linear differential equations. In equation (73) x is a vector of system states, $\dot{x}$ is the derivative of the system states, F is the system dynamics matrix, G is the input gain matrix, u is a known control vector and w is a white (Gaussian) noise vector.

$$\dot{x} = Fx + Gu + w \tag{73}$$

Equation (73) is called the state space or matrix form of the differential equations. The quintessential example of a state space model is the position, velocity, and acceleration of an object that you are tracking. The process noise matrix Q is defined by the process noise vector.

$$Q = E[ww^T] \tag{74}$$

Another assumption when using the Kalman filter is that measurements are linearly related to the system states. Tracking an object using range and bearing information converted into a Cartesian coordinate system is an example where measurements are not linearly related to system states.

$$z = Hx + v \tag{75}$$

In equation (75) z represents the measurement vector, H is the measurement matrix, x is the true system state and v is the Gaussian measurement noise vector.

$$R = E[vv^T] \tag{76}$$

In a manner similar to the process noise matrix, the measurement noise matrix is defined by the measurement noise. A transition or fundamental matrix Φ can be used to exactly propagate forward the state of any linear time invariant (LTI) system from time t_0 to t [Zarchan 2000]. The time step is defined as $T_s = t - t_0$. The exact method for determining the fundamental matrix is shown in equation (77) and the Taylor series approximation is shown in (78).

$$\Phi(t) = L^{-1}[(sI - F)^{-1}] \tag{77}$$

$$\Phi(t) = e^{Ft} = I + Ft + \frac{(Ft)^2}{2!} + \cdots + \frac{(Ft)^n}{n!} \tag{78}$$

Digitally sampled systems are by their nature discrete time systems. The discrete fundamental matrix, at time step k, is calculated by substituting the time step for t as in equation (79). Similarly the discrete measurement vector, and measurement matrix are calculated in equations (80) and (81).

$$\Phi_k = \Phi(T_s) \tag{79}$$

$$z_k = Hx_k + v_k \tag{80}$$

$$R_k = E[v_k v_k^T] \tag{81}$$

Conveniently in polynomial Kalman filters the measurement noise matrix is a scalar. The discrete process noise matrix is evaluated as an integral shown in equation (82).

$$Q_k = \int_0^{T_s} \Phi(\tau) Q \Phi_k^T d\tau \tag{82}$$

The one step ahead projection $\bar{x}$ based on the system dynamics and any known inputs to the system is calculated in equation (83). This term represents the projected state of the system one time step into the future. This projection should not be confused with the multiple time step projections described later in the document when the state estimation algorithms are used as part of a prognostic algorithm.

$$\bar{\mathrm{x}} = H\Phi_k \hat{x}_{k-1} - HGu_{k-1} \tag{83}$$

Finally the Kalman filter equation is presented.

$$\hat{x}_k = \Phi_k \hat{x}_{k-1} + G_k u_{k-1} + K_k(z_k - \bar{\mathrm{x}}) \tag{84}$$

Where K_k is the Kalman gain and the discrete input gain matrix is calculated as,

$$G_k = \int_0^{T_s} \Phi(\tau) G \, d\tau \tag{85}$$

It is assumed that u_{k-1} is constant between time steps. It is difficult to draw comparisons of the most general Kalman filter equation to the recursive least squares update equation (50). If a number of constraints are to the equation (82), as in the case of the scalar Kalman filter presented in the next section, the similarities will become more obvious. It can be shown that the matrix Riccati equations can be used to solve for the Kalman gain.

$$M_k = \Phi_k P_{k-1} \Phi_k^T + Q_k \tag{86}$$

$$K_k = M_k H^T (H M_k H^T + R_k)^{-1} \tag{87}$$

$$P_k = (I - K_k H) M_k \tag{88}$$

Where M_k is the covariance matrix representing errors in the state estimates prior to a measurement update, but after calculating the one step ahead projection. P_k is the covariance

matrix representing errors in the state estimates after a measurement update. To initialize the recursive Kalman filter you must define $\hat{x}_0, P_0, R_0, and\ Q_0$. You must determine before running the filter the differential equations that describe your system, and cast them in state or matrix form. To build intuition about the Kalman filter and Riccati equations the simple case of a scalar Kalman filter will be discussed next.

5.2.2 Derivation of Scalar Ricatti Equations

The scalar Kalman filter is a simple case where the matrix equations degenerate into scalar equations and provide a useful insight into the operation of the filter. If there are no inputs into the system the discrete system model reduces to a scalar equation (89). The only state in this filter is the state being measured. Higher order filters allowing the estimation of derivatives and model parameters was described in the previous section and will be demonstrated in practice later in the document during the baseline code validation examples.

$$x_k = \Phi_k x_{k-1} + w_k \tag{89}$$

The one step ahead projection is also simplified under the no input constraint.

$$\bar{\mathrm{x}} = H\Phi_k \hat{x}_{k-1} \tag{90}$$

The Kalman update equation (91) is now in a form that is much more similar to the recursive least squares update equation shown in (50). In the Kalman filter the residual is between the measurement and the projected state of the system, $z_k - \bar{\mathrm{x}}$. Recursive least squares filtering uses previous state values for updating, while the Kalman filter uses projected state values for updating system states. The most important distinction is how the gain term is calculated in the Kalman filter. It will be shown that under the restrictive set of constraints implied by applying the Kalman filter equations that the gain is the optimal gain for minimizing the variance in state estimates. Later sections will address the relaxation of strict linear time

invariant system dynamics and Gaussian noise models with the extended Kalman filter and particle filter.

$$\hat{x}_k = \Phi_k \hat{x}_{k-1} + K_k(z_k - \bar{\mathrm{x}}) \tag{91}$$

In the scalar Kalman filter the measurement vector, measurement matrix, true state and noise vector reduce to a scalars.

$$\mathrm{z_k} = Hx_k + v_k \tag{92}$$

The error in the estimate $\tilde{x}_k$ is the difference between the true state and the estimated state.

$$\tilde{x}_k = x_k - \hat{x}_k = x_k - \Phi_k \hat{x}_{k-1} + K_k(z_k - \bar{\mathrm{x}}) \tag{93}$$

By using equations (89), (92), and (95) the error can be written in terms of only the state estimate, measurement scalar, Kalman gain and noise terms.

$$\mathrm{z_k} = Hx_k + v_k = H(\Phi_k x_{k-1} + w_k) + v_k \tag{94}$$

$$\tilde{x}_k = x_k - \hat{x}_k \rightarrow \tilde{x}_{k-1} = x_{k-1} - \hat{x}_{k-1} \tag{95}$$

$$\tilde{x}_k = (1 - K_k H)\hat{x}_{k-1} + (1 - K_k H)w_k - K_k v_k \tag{96}$$

The covariance scalar representing errors in the state estimates after a measurement update is defined as the expectation of the square of the estimation error.

$$P_k = E[\tilde{x}_k^2] = (1 - K_k H)^2(P_{k-1}\Phi_k^2 + Q_k) + K_k^2 R_k \tag{97}$$

The scalar versions of the process noise and measurement noise are.

$$Q_k = E[w_k^2] \tag{98}$$

$$R_k = E[v_k^2] \tag{99}$$

To simplify the error covariance matrix define M_k. This is the scalar equivalent of the first Ricatti equation originally presented in equation (86).

$$M_k = P_{k-1}\Phi_k^2 + Q_k \tag{100}$$

Substituting (100) into (97) and expanding gives.

$$P_k = (1 - K_k H)^2 M_k + K_k^2 R_k = M_k - 2K_k H M_k + K_k^2 H^2 M_k + K_k^2 R_k \tag{101}$$

To find the Kalman gain that minimizes the variance of the state estimate, set the derivative of (101) equal to zero.

$$\frac{\partial P_k}{\partial K_k} = 0 = 2(1 - K_k H)M_k(-H) + 2K_k R_k \tag{102}$$

Solving (102) for K_k provides a scalar expression that is equivalent to the second Ricatti equation originally presented in (87).

$$K_k = \frac{M_k H}{H^2 M_k + R_k} \tag{103}$$

To derive the scalar version of the final Riccati equation re-arrange (103) and substitute into (101). By simplifying the result you obtain the equivalent formulation of the third Ricatti equation in equation (107).

$$K_k R_k = M_k H - H^2 M K_k \tag{104}$$

$$P_k = M_k - 2K_k H M_k + K_k^2 H^2 M_k + K_k(M_k H - H^2 M K_k) \tag{105}$$

$$P_k = M_k - 2K_k H M_k + K_k^2 H^2 M_k + K_k M_k H - K_k^2 H^2 M \tag{106}$$

$$P_k = (1 - K_k H)M_k \tag{107}$$

It is insightful to investigate the influence of the process noise and measurement noise terms on the scalar Ricatti equations. There are best practices for choosing the process and measurement noise terms, but ultimately these terms act as free parameters and can drastically influence the performance of the filter. For most applications selecting the parameters within the correct order of magnitude is sufficient. In other scenarios it may be desirable to add fictitious

noise to the filter to achieve better smoothing performance. The scalar Kalman gain equation is repeated here for convenience. Note that the process noise term directly effects the M_k term.

$$K_k = \frac{M_k H}{H^2 M_k + R_k} \tag{108}$$

- Case 1: if M_k is much smaller than 1, then K_k is much smaller than 1, therefore disregard the measurement
- Case 2: if M_k is much greater than R_k, then $K_k = 1/H$, therefore disregard the estimate
- Case 3: if R_k is much greater than M_k, then K_k is much smaller than 1, therefore disregard measurement

5.2.3 Extended Kalman Filter

Describing processes with linear differential equations and measurement equations that are linear functions of the underlying states may not always accurately model real world phenomena. The Extended Kalman filter allows a system to be described by a set of non-linear differential equations and non-linear measurement equations. The extended Kalman filter applies a first order linearization in the Riccati equations. The more general particle filter discussed later does not require linearization.

A set of higher order differential equations can be cast as a set of first order differential equations through the repeated substitution of intermediate variables [Kreyszig 2010]. To implement an Extended Kalman filter, a system of nonlinear first order differential equations is used to describe the system.

$$\dot{x} = f(x) + w \tag{109}$$

Where x is the vector of system sates, $f(x)$ is a non-linear function of the system states, and w is still a Gaussian noise term. The measurement equation is no longer limited to linear equations.

$$\mathrm{z} = h(x) + v \tag{110}$$

Where $h(x)$ is a nonlinear function of the measured states and v is still a Gaussian noise term. An example of a nonlinear measurement equation is the situation where navigation measurements are reported in bearing and range format (i.e. Polar coordinates) and must be transformed into Cartesian coordinates through the use of nonlinear trigonometric relationships. To calculate the system dynamics matrix F and measurment matrix H required in the matrix Riccati equations a first order approximation is used to represent the nonlinear functions $f(x)$ and $h(x)$.

$$\mathrm{F} = \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x}} \tag{111}$$

$$\mathrm{H} = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}} \tag{112}$$

The logic used earlier in the Kalman filter section to describe the discrete versions of the process noise, measurement noise, and fundamental matrix is reproduced here for convenience.

$$\mathrm{R_k} = E[v_k v_k^T] \tag{113}$$

$$Q = E[ww^T] \tag{114}$$

$$Q_k = \int_0^{T_s} \Phi(\tau) Q \Phi_k^T d\tau \tag{115}$$

$$\Phi_\mathrm{k} = \Phi(T_s) = e^{FT_s} = I + FT_s + \frac{(FT_s)^2}{2!} + \cdots + \frac{(FT_s)^n}{n!} \tag{116}$$

The discrete form of the measurement equation incorporates the nonlinear measurement function

$$\mathrm{z_k} = h(x_k) + v_k \tag{117}$$

The extended Kalman filter equation relates the projected one step ahead state of the system with a residual and the Kalman gain term. The Kalman gain is calculated with the matrix Riccati equations.

$$\hat{x}_k = \bar{x}_k + K_k(z_k - \mathrm{h}(\bar{x}_k)) \tag{118}$$

There are no constraints on how the one step ahead projection is calculated. In this work Euler integration will be used to calculate the projection.

Where T_E is the Euler integration time step. If appropriate T_E and T_s may be equal, but they do not have to be. Due to the linear approximation in the Riccati, the extended Kalman filter is dependent on the appropriateness of the linear approximation over the length of the time step.

5.2.4 Bayesian Filtering

The particle filter represents the state of the art in non-linear filtering and is a generalization of Kalman and extended Kalman filtering methods. Particle filtering follows the same general steps as the Kalman family filters, but utilizes a Bayesian framework for calculating state estimates as probabilities. The risk based decision making framework used in this document is also based on Bayesian statistics, so the two methods dovetail nicely.

5.2.4.1 Bayesian Statistics and Baye's Theorem

Frequentist statistics are the class of statistics used widely in traditional reliability approaches. In the frequentist framework, a probability is defined as the limiting value of an event occurring as a number of observations approaches infinity. It can be seen from the definition that there are a number of problems that can arise when applying frequentist statistics to ensure the correct operation of safety and mission critical electronic systems. For example the frequentist approach requires a large number of observations and generalizes reliability in terms

of large population sizes. Safety and mission critical electronics require probabilities of failure defined for a single component. Further it is prohibitively expensive to generate a large number of observations for large complex engineering systems. The consequences of a failure in a many systems (e.g. nuclear power generation, jet engines) may make the acquisition of run to failure data impossible to obtain.

Bayesian statistics state that a probability is the belief that a variable is a certain value. This approach has many advantages in its formulation when applied to ensuring reliability of safety and mission critical electronics. Inherently Bayesian statistics are applicable to a single system. Frequentists readily point out that Bayesian statistics are subjective because they depend on your belief in the state of the system prior to performing inference. It is acceptable in some cases to assume initially that you are ignorant about the state of your system and avoid the need for expensive run to failure data.

Bayesian statistics depend heavily on the use of Baye's theorem. The theorem is easy to derive from first principles of probability. First the sum rule and product rule for probabilities are stated. Then the two rules are manipulated into Baye's theorem.

The sum rule is stated as the probability of an event x occurring, plus the probability of the same event x not occurring is defined as one. The tilde symbol represents the NOT operator.

$$p(x) + p(\sim x) = 1 \tag{119}$$

The product rule is stated as the probability that two events x and y will both occur and is equal to the conditional probability of x occurring if y occurs times the probability that y occurs.

$$p(x, y) = p(x|y)p(y) \tag{120}$$

Where the comma is the AND operator and the bar "|" is the conditional operator. The probability $p(x|y)$ reads as the probability that x occurs given that y occurred.

To derive Baye's theorem transpose the events in equation (120).

$$p(y,x) = p(y|x)p(x) \tag{121}$$

Since $p(y,x) = p(x,y)$

$$p(y|x)p(x) = p(x|y)p(y) \tag{122}$$

Rearranging the previous equation results in Baye's theorem.

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \tag{123}$$

In terms of the scientific process, the theorem can be stated as the probability of a hypothesis being true given observed data

$$p(hypothesis|data) = \frac{p(data|hypothesis)p(hypothesis)}{p(data)} \tag{124}$$

The value in the denominator is not easily observed in practice in the form of a marginal probability. Using the law of total probability relates the marginal probability to its conditional probabilities.

$$p(y) = \sum_n p(y|x_n)p(x_n) \tag{125}$$

For the case of a binary variable, the law of total probabilities reduces to

$$p(y) = p(y|x)p(x) + p(y|{\sim}x)p({\sim}x) \tag{126}$$

And equation (124) becomes

$$p(hypothesis|data) = \frac{p(data|hypothesis)p(hypothesis)}{p(hypothesis|data)p(data) + p(hypothesis|\sim data)p(\sim data)} \quad (127)$$

To illustrate this point consider a hypothetical example of a crack in an airplane wing. From historical records it is known that after 1000 flight hours there is a 1% probability of a crack existing in a wing. Traditional inspection methods can only detect cracks after they grow to a dangerous length. A new experimental method has been introduced to aid in detecting cracks before traditional methods to enable extended use out of the wing and thereby reduce operating costs. Under the traditional method wings are replaced after 1000 flight hours. The method correctly identifies wings with cracks before any other method is capable of detecting a crack 80% of the time. Unfortunately there is a 9.6% false positive rate associated with the technique. If a wing is flagged by the new system as having a crack, what is the probability that the wing actually has a crack?

Using Bayes theorem as stated in (127), the following probabilities are extracted from the hypothetical problem. The hypothesis is a crack in the wing. The observed data is the positive flag from the new test method. It is desired to know $p(crack|positive)$.

$$p(hypothesis) = p(crack) = 0.01 \quad (128)$$

$$p(data|hypothesis) = p(positive|crack) = 0.8 \quad (129)$$

$$p(data|\sim hypothesis) = p(positive|\sim crack) = 0.096 \quad (130)$$

$$\begin{aligned} p(hypothesis|data)p(data) & \\ + p(hypothesis|\sim data)p(\sim data) & \\ = p(positive|crack)p(crack) & \\ + p(positive|\sim crack)p(\sim crack) & \\ = (0.8)(0.01) + (0.096)(0.01) = 0.103 & \end{aligned} \quad (131)$$

$$\begin{aligned} p(hypothesis|data) = p(crack|positive) &= \frac{(0.8)(0.01)}{0.103} \\ &= 0.078 = 7.8\% \end{aligned} \quad (132)$$

Given the hypothetical inputs, and using Baye's theorem provides us with the inference that there is a 7.8% probability of a crack existing in the wing. In addition to demonstrating the use of Baye's theorem, this example highlights the disastrous effect of uncertainty that false positive predictions can have on an inference. Even with a hypothetical method that gives a perfect prediction of a crack if it exists, the inference can still be quite low.

Table 6: Inference with a perfect test and different false positive rates

True Positive Rate: $p(positive\|crack)$	False Positive Rate: $p(positive\|\sim crack)$	Inference: $p(crack\|positive)$
100%	9.6%	9.5%
100%	1.0%	50.3
100%	0.1%	91%

Bayesian statistics are considered subjective because your initial belief in the state of the system affects your results. Table 7 shows the results of taking different initial beliefs that a crack exists in the wing at 1000 hours.

Table 7: Inference with a perfect test and different initial beliefs in the system state

True Positive Rate: $p(positive\|crack)$	False Positive Rate: $p(positive\|\sim crack)$	Initial Belief: $p(crack)$	Inference: $p(crack\|positive)$
100%	9.6%	1%	9.5%
100%	9.6%	10%	53.6
100%	9.6%	50%	91.2%

5.2.4.2 Bayesian Framework for State Estimation

The Bayesian framework for state estimation is derived from first principles described in the previous section. Restating the product rule with new variables. The 'and' operator has been omitted for brevity.

$$p(AB) = p(BA) = p(A|B)p(B) = p(B|A)p(A) \tag{133}$$

Baye's theorem for two events is stated as

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \tag{134}$$

Baye's theorem for three events is derived by substituting B = BC into the two event formula and then repeatedly applying the product rule.

$$p(A|BC) = \frac{p(BC|A)p(A)}{p(BC)} = \frac{p(B|AC)p(A|C)}{p(B|C)} \tag{135}$$

A discrete form of the Champman-Kolomgoron equation can be derived starting with the product rule.

$$p(A|B) = \frac{p(AB)}{p(B)} = \frac{\sum_c p(ABC)}{p(B)} = \sum_c p(A|CB)p(C|B) \tag{136}$$

For a continuous probability density function, equation (136) can be rewritten as

$$p(A|B) = \int p(A|CB)p(C|B)dA \tag{137}$$

In the Bayesian state estimation framework future states x_k are a function of prior states x_{k-1} and a noise term w_{k-1}.

$$x_k = f_{k-1}(x_{k-1}, w_{k-1}) \tag{138}$$

Similarly the measurement equation is a function of the system states and a separate noise term.

$$z_k = h_k(x_k, v_k) \tag{139}$$

Lower case z_k represents a single measurement. Uppercase Z_k represents the vector of all previous observations such that $Z_k = \{z_i, i = 1 \dots, k\}$. Filtering follows the same steps as the Kalman family of filters, except states are represented as generic PDF's.

5.2.4.2.1 One Step Ahead Prediction

The future state of the system predicted one time step into the future using the prior knowledge about the system is calculated using (137). Assuming that $p(x_{k-1}|Z_{k-1})$ is available,

$$p(x_k|Z_{k-1}) = \int p(x_k|x_{k-1}, Z_{k-1})p(x_{k-1}|Z_{k-1})dx_{k-1} \tag{140}$$

Future system states are not dependent on prior measurements, so

$$p(x_k|x_{k-1}, Z_{k-1}) = p(x_k|x_{k-1}) \tag{141}$$

The one step ahead prediction $\bar{x}_k$ is therefore

$$\bar{x}_k = p(x_k|Z_{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|Z_{k-1})dx_{k-1} \tag{142}$$

The first term in the one step ahead prediction integral is calculated from equation (138), and the second term inside the integral is assumed to be known.

5.2.4.2.2 Measurement Step

The measurement step is performed according to equation (139), and Z_k has now been observed. There are no restrictions on the equations in the measurement equation. The equations can even be represented by non-continuous PDF's. For example if there is known to be a wall in the environment a robot is operating in this can be reflected in the measurement equation in the form of a non-continuous PDF.

5.2.4.2.3 State Estimate Update

The estimated state of the system $\hat{x}_k = p(x_k|Z_k)$ is desired by updating the belief in the one step ahead projection using the newly available measurement.

$$p(x_k|Z_k) = p(x_k|z_k, Z_{k-1}) \tag{143}$$

Using the two event Baye's theorem from equation (135).

$$p(x_k|Z_k) = \frac{p(z_k|x_k, Z_{k-1})p(x_k|Z_{k-1})}{p(z_k|Z_{k-1})} \tag{144}$$

The fact that future measurements are not dependent on past measurements was used to simplify the previous equation.

$$\hat{x}_k = p(x_k|Z_k) = \frac{p(z_k|x_k)p(x_k|Z_{k-1})}{p(z_k|Z_{k-1})} \tag{145}$$

Finally the update equation can be written. Note that the denominator, often denoted as the normalizing constant can be expanded using (137).

$$p(z_k|Z_{k-1}) = \int p(z_k|x_k)p(x_k|Z_{k-1})dx_k \tag{146}$$

In the integral in equation (146) the first term is calculated from (139). In theory the second term is calculated from equation (142), but in practice equation (146) is an intractable integral that is not solvable. The implementation of the particle filter that will be presented in the next section is designed to evaluate the predict-measure-update equations presented in this section without evaluating the integral in the denominator of the update equation.

5.2.4.3 Particle Filter

The particle filter represents the probability distribution over the current value of each state variable using a discrete probability mass function. Unlike with Kalman filtering there are no restrictions on the shape of the distribution function being approximated. The N discrete

weighted samples of the probability density function are called particles and located in the states space at x_k^i. Probability of each particle is denoted by its weight, w_k^i.

$$p(x_k|Z_k) \approx \sum_{i=1}^{N} w_k^i \delta(x_k - x_k^i) \tag{147}$$

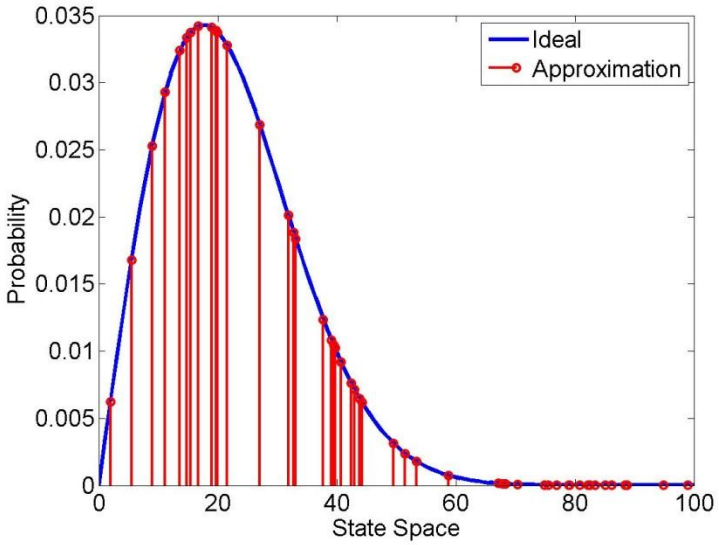


Figure 41: Approximation of an ideal and continuous probability density function by a set of particles. Particles locations represented as $\boldsymbol{x_k^i}$ and specify the position on the x-axis. Weights specify the height of the stems and are represented by $\boldsymbol{w_k^i}$

As the number of particles increases to infinity the approximation converges on the underlying continuous probability density function. In general particles will not be evenly spaced. Particles far from the regions with the highest probability will have a negligibly low weight. The Driac delta function δ is defined such that

$$\delta(x-a) = 0 for\ x \neq a\ ;\ \int_{-\infty}^{\infty} \delta(x-a)dx = 1 \tag{148}$$

Weights are normalized to represent a proper probability mass function

$$\sum_{i=1}^{N} w_k^i = 1 \tag{149}$$

The particle filtering algorithm is summarized below. Notice how the process is identical to the Kalman family of filters, but the details of the process now utilize Bayesian statistics.

SIR Particle Filtering Algorithm

1. Initialize a distribution of particles, $\{x_0^i, w_0^i\}\ where\ i = 1,2, \dots, N$
2. Project the particles forward one time step to find $\bar{x}_k^i = p(x_k|Z_{k-1})$
3. Make a new system measurement, z_k
4. Assign new weights based on the relevance of each particle compared to the new measurement $\hat{x}_k^i = p(x_k|Z_k)$
5. Resample: If a few particles contain a majority of all possible weights, kill low probability particles and replace them with new more relevant particles

5.2.4.3.1 Initialize a distribution of particles

The particle filter is initialized with a probability mass function to represent the initial state guess of all the state variables. Common choices are to draw from a normal distribution with mean μ and standard deviation σ, or alternatively to draw from a unifom distribution $U(a, b)$ with possible values defined on the interval [a,b].

5.2.4.3.2 One step ahead projection

There are no restriction on how to propagate the particles for the one step ahead prediction. Euler integration will be used in this document, but any model that can propagate the discrete sets of particles in time will suffice.

$$\bar{x}_k^i = \hat{x}_{k-1}^i + \frac{\partial \hat{x}_k^i}{\partial t} T_E + w \tag{150}$$

The addition of the process noise term, $w \sim N(0, \sigma_w)$, provides enough variation in the particles to account for un-modeled effects.

5.2.4.3.3 System Measurement

Any measurement methodology is acceptable for obtaining z_k in the particle filtering framework.

5.2.4.3.4 Update state estimate and assign new weights

When a new measurement z_k becomes available at time k, a penalty function is used to describe $(z_k|x_k^i)$ and re-assign weights based on the relevance of each particle compared to the new measurement. Choice of weights w_k^i, are chosen based on the principle of importance sampling. If the particles x_k^i are drawn from an importance density $q(x_k^i|z_k)$, then the weights are represented by:

$$w_k^i \, \alpha \, \frac{p(x_k^i|z_k)}{q(x_k^i|z_k)} = w_{k-1}^i \frac{p(z_k|x_k^i)p(x_k^i|x_{k-1}^i, Z_k)}{q(x_k^i|X_k^i, Z_k)} \tag{151}$$

Where $X_k^i = \{x_n^i, n = 1 \ldots, k\}$. The optimal importance density function that minimizes the variance of importance weights, conditioned upon x_k and z_k has been shown to be to take the importance density to be equal to the prior probability [Ristic 2004]:

$$q(x_k^i|X_k^i, Z_k) \approx p(x_k^i|x_{k-1}^i, Z_k) \tag{152}$$

This results in the update equation

$$w_k^i = w_{k-1}^i p(z_k|x_k^i) \tag{153}$$

The particle filter is not limited to any particular penalty function, so Gaussian kernel function was used for re-assigning weights.

$$w_k^i = w_{k-1}^i \frac{1}{\sqrt{2\pi\sigma_v}} e^{\frac{-(z_k - x_k^i)}{(2\pi\sigma_v)}} \tag{154}$$

It is important after re-assignment to normalize the weights to maintain the approximated probability mass function as a proper distribution where the sum of the weights equals unity.

$$w_k^i = \frac{w_k^i}{(\sum_{i=1}^N w_k^i)} \tag{155}$$

5.2.4.3.5 Resample as necessary

If a few particles contain a majority of all of the possible weights, kill the low probability particles and replace them with more relevant particles moved to higher probability locations. This problem is called particle degeneracy and a simplistic case is shown in Figure 42. If degeneracy is not severe, this step is skipped until needed. Resampling is required if

$$\frac{1}{\sum\left(w_k^i\right)^2} < \frac{N}{4} \tag{156}$$

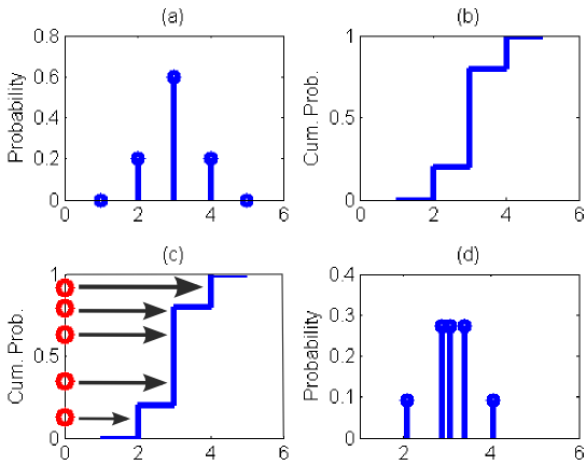


Figure 42: a) initial probability mass function represented by five particles, b) cumulative distribution function, c) assigning draws from uniform distribution U(0,1) to new particle locations based on cumulative distribution function, d) importance re-sampled distribution

For a complete discussion beyond the brief intuitive overview of particle filtering provided here see [Ristic 2004, Cappé 2007].

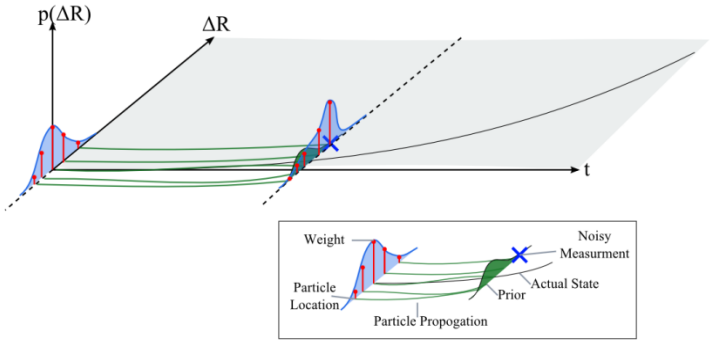


Figure 43: Evolution of Model Prediction in Particle Filter

5.3 Tracking Example: Physics of a Ballistic Object Re-entry Path

To facilitate a fair benchmark test for the recursive algorithms implemented in this document, a widely used synthetic data set [Farina 2002, Srinivasan 2006, Ristic 2004, Zarchan 2000] describing the re-entry of a ballistic object into the earth's atmosphere will be studied. By benchmarking performance against a simulated event the true, noise free, state is available for judging performance of the implemented filter. In practice the true underlying state of the system can never be fully observed.

The flight of a ballistic object can be categorized into three phases [Ristic 2004]. Phase one is a powered launch, where both gravity and air drag act on the object. Phase two, is an un-powered flight above the earth's atmosphere where only the force of gravity is significant. Phase three, which will be simulated for the benchmark test, is the un-powered re-entry of the ballistic object into the earth's atmosphere under the influence of both gravity and air drag. Due to the gradually increasing influence of drag as the atmosphere becomes denser with decreasing elevation, the dynamic models describing the third and final stage of a ballistic objects flight are usually non-linear. A simplistic one dimensional non-linear dynamic model with additive

Gaussian noise is used to create the synthetic benchmark data set. Besides the obvious applications of this data set to missiles and defense applications, the re-entry of space debris [Welford 2008] can be modeled and tracked in a similar manner.

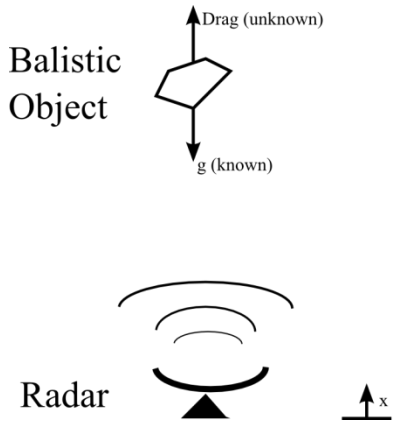


Figure 44: Simple example problem for generating a baseline data set

Drag, denoted as D, is a function of the air density ρ, the velocity of the object, and a ballistic coefficient parameter, β. Air density is itself a function of height x.

$$D = \frac{\rho(x) g \dot{x}^2}{2\beta} \tag{157}$$

$$\rho(x) = \gamma e^{-\eta x} \tag{158}$$

The position, velocity, and acceleration of the system can be written in state space form as

$$x_k = \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} \tag{159}$$

$$\begin{aligned} x &= x - \tau\dot{x} \\ \dot{x} &= \dot{x} + (D - g)\tau \\ \ddot{x} &= \ddot{x} \end{aligned} \tag{160}$$

Or in the more useful matrix form,

$$f(x_k) = \Phi x_k - G(D(x_k) - g) \tag{161}$$
$$= \begin{bmatrix} 1 & -\tau & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} + \begin{Bmatrix} 0 \\ \tau \\ 0 \end{Bmatrix} (D(x_k) - g)$$

Using a popular set of input parameters ($\tau = 0.005\ sec, \gamma = 1.754 \frac{kg}{m^3}, \eta = 1.49x10^{-4}, g = 9.81 \frac{m}{s^2}, x_0 = 61000\ m, \dot{x}_0 = 3048 \frac{m}{s}, \ddot{x}_0 = 0 \frac{m}{s^2}, \beta = 19161$) the following ideal trajectory was simulated.

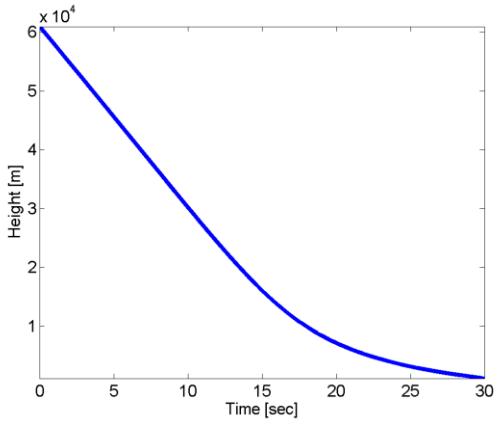


Figure 45: Simulated position of ballistic object

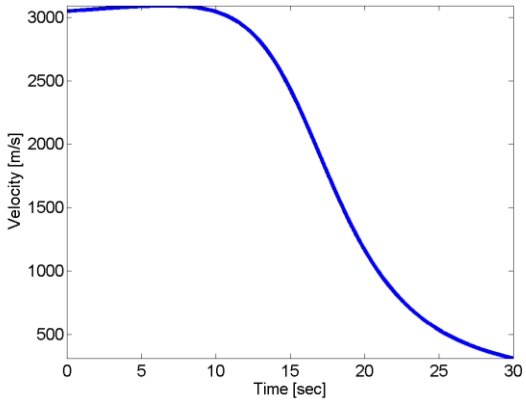


Figure 46: Simulated velocity of ballistic object

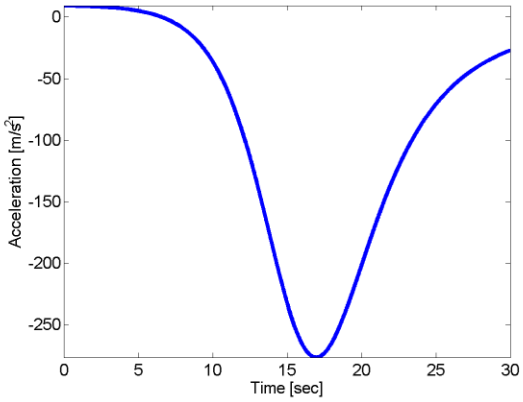


Figure 47: Simulated acceleration of ballistic object

The effect of the ballistic coefficient, β, can be quite dramatic on the trajectory of the object.

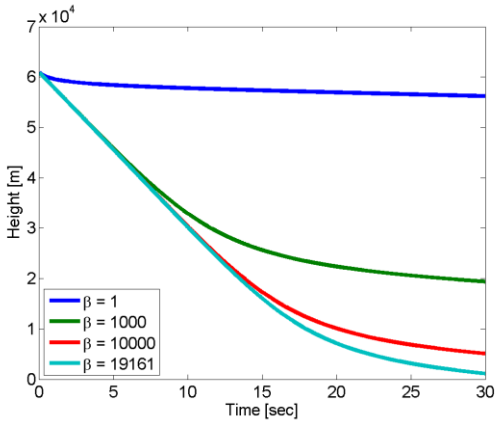


Figure 48: Effect of drag coefficient on simulation

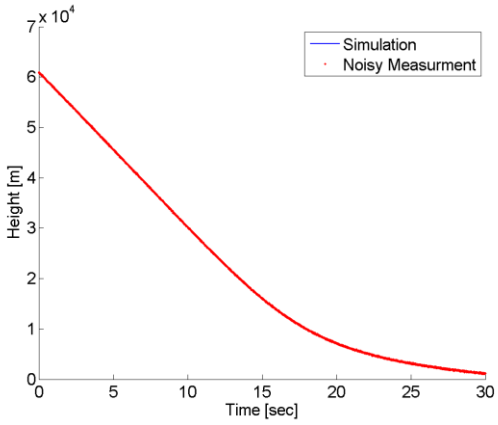


Figure 49: Simulating a noisy measurement from the ground truth simulation. The simulation line is essentially the same, and therefore not visible in the plot, as the noisy measurement at this level of resolution

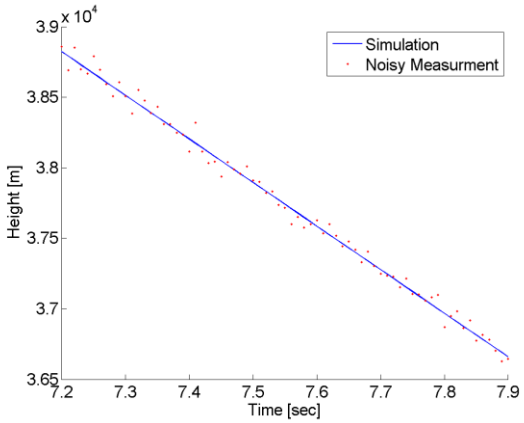


Figure 50: Zoomed view of the noisy measurement

If a noisy measurement is taken with a standard deviation of 7.62 meters (25 ft.), then the following noisy measurement is available for tracking. Viewed from the 60000 meter level, the measurement is quite accurate. A zoomed view of the same noisy measurement is provided to put the scale of the variation in perspective.

The remainder of the document will use the noisy measurement data set as the input into all of the different algorithms. A variety of cases will be investigated involving the Kalman filter and extended Kalman filter. The case of a known ballistic coefficient, an unknown ballistic coefficient, and finally the case of a time evolving ballistic coefficient will be demonstrated to compare and contrast the algorithms.

5.3.1 Tracking a known object (β is known and constant)

For the case of a known, constant ballistic constant two filters are compared. First the Kalman filter, and then the extended Kalman filter. Based on the known noise level in the radar measurement the variance for both filters is taken to be $R = (7.62\,m)^2 = 58.06\,m^2$. The maximum change in acceleration during tracking is estimated to be 10g's. Therefore the process

noise term is estimated as the variance of the expected acceleration divided by the length of the tracking time, $\Phi_s = \frac{(10g)^2}{30} = 320$.

5.3.1.1 Kalman Filter

The state vector for the Kalman filter is taken as

$$x_k = \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} \tag{162}$$

In this case the derivative of acceleration is taken to be white noise with a standard deviation of u_s. The state space equation becomes

$$\begin{Bmatrix} \dot{x} \\ \ddot{x} \\ \dddot{x} \end{Bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ u_s \end{Bmatrix} \tag{163}$$

By inspection,

$$F = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{164}$$

$$\Phi = e^{Ft} \approx I + Ft + \frac{(Ft)^2}{2!} + \cdots \tag{165}$$

$$\Phi_{1\ term} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} t = \begin{bmatrix} 1 & -t & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \tag{166}$$

$$\Phi_{2\ term} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} t + \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} t^2 = \begin{bmatrix} 1 & -t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \tag{167}$$

Since the Φ matrix is used to propagate the state estimate, it would be desirable to extend the Taylor series approximation to many terms. Unfortunately the F^2 term and above are

identically zero, so the two term approximation is the upper limit of the accuracy of the approximation.

The discrete process noise matrix is formed from the continuous process noise matrix as

$$Q = \Phi_s \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; Q_k = \int_0^{Ts} \Phi(\tau) Q \Phi(\tau)^T d\tau \tag{168}$$

$$Q_{k,1\ term} = \Phi_s \begin{bmatrix} 0 & 0 & 0 \\ 0 & \frac{Ts^3}{3} & \frac{Ts^2}{2} \\ 0 & \frac{Ts^2}{2} & Ts \end{bmatrix} ; \ Q_{k,2\ term} = \Phi_s \begin{bmatrix} \frac{Ts^5}{20} & \frac{Ts^4}{8} & \frac{Ts^3}{6} \\ \frac{Ts^4}{8} & \frac{Ts^3}{3} & \frac{Ts^2}{2} \\ \frac{Ts^3}{6} & \frac{Ts^2}{2} & Ts \end{bmatrix} \tag{169}$$

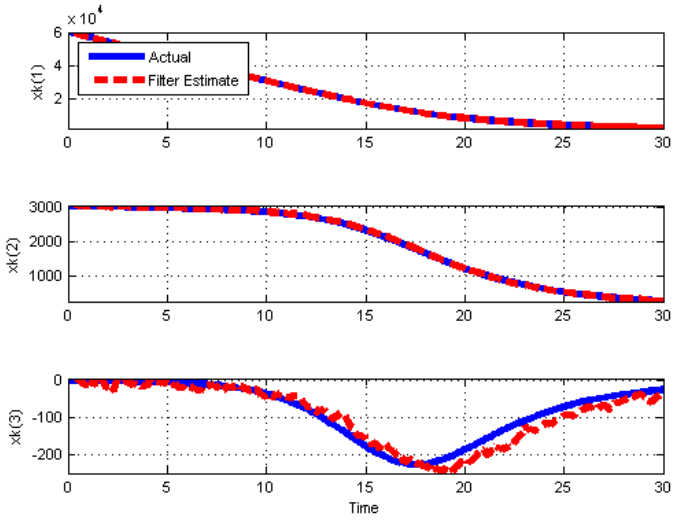


Figure 51: Tracking results for a two term PVA Kalman filter

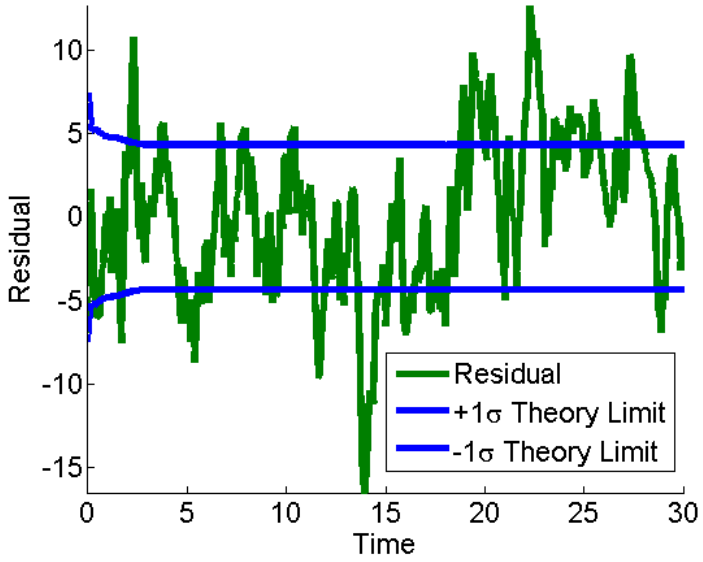


Figure 52: Residual of the error between the estimated and simulated (no noise) states

Results from the implemented Kalman filter are shown in Figure 51. As the influence of the non-linear drag term becomes numerically significant around 10 seconds into the simulation the tracking on the acceleration state starts to degrade. The position and velocity states are estimated with good accuracy even if assumptions about linearity in the system are being violated. Figure 52 shows a plot of the residual between the estimated state and the simulated actual position (no noise). When the Kalman filter is implemented in an optimal manner, the residuals fall inside the $\pm\sigma$ bounds 63% of the time [Zarchan 2000].

5.3.1.2 Extended Kalman Filter

The state vector for the Extended Kalman filter is taken as

$$x_k = \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} \tag{170}$$

In this case the derivative of acceleration is taken to be white noise with a standard deviation of u_s. The state space equation becomes

$$f(x_k) = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ u_s \end{Bmatrix} \tag{171}$$

The fundamental matrix is now determined using the Jacobian of the state vector

$$F = \frac{\partial f}{\partial x_k}\Big|_{x_k} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{172}$$

This result is the same as the Kalman filter, which is not surprising since nothing about the dynamics of the system has been assumed to be different than the case of the Kalman filter. The next EKF development will use a more detailed acceleration model which will change the implementation of the filter.

5.3.1.3 Extended Kalman Filter with Non-Linear Acceleration

The state vector for the Extended Kalman filter is taken as

$$x_k = \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} \tag{173}$$

In this case the derivative of acceleration is taken to be white noise with a standard deviation of u_s. The state space equation becomes

$$f(x_k) = Ax_k - G[D(x_k) - g] \tag{174}$$

Where

$$\mathrm{A} = \begin{bmatrix} 1 & -\tau & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} ; D(x_k) = \frac{\rho(x) g \dot{x}^2}{2\beta} ; \rho = \gamma e^{-\eta x} \tag{175}$$

Equivalently this the same as

$$\begin{aligned} x &= x - \tau\dot{x} \\ \dot{x} &= \dot{x} - \tau(D(x_k) - g) \\ \ddot{x} &= \ddot{x} \end{aligned} \tag{176}$$

Which means that the differential equations for the system are

$$\begin{aligned} \dot{x} &= & -\dot{x} \\ \ddot{x} &= -(D(x_k) - g) &= \frac{-\gamma e^{-\eta x} g \dot{x}^2}{2\beta} + g \\ \dddot{x} &= & u_s \sim N(0, \Phi_s) \end{aligned} \tag{177}$$

The terms for the Jacobian become

$$\partial \dot{x}/_{\partial x} = f_{11} = 0 \tag{178}$$

$$\partial \dot{x}/_{\partial \dot{x}} = f_{12} = -1 \tag{179}$$

$$\partial \dot{x}/_{\partial \ddot{x}} = f_{13} = 0 \tag{180}$$

$$\partial \ddot{x}/_{\partial x} = f_{21} = \frac{\eta\gamma e^{-\eta x} g \dot{x}^2}{2\beta} \tag{181}$$

$$\partial \ddot{x}/_{\partial \dot{x}} = f_{22} = \frac{-\gamma e^{-\eta x} g \dot{x}}{\beta} \tag{182}$$

$$\partial \ddot{x}/_{\partial \ddot{x}} = f_{23} = 1 \tag{183}$$

$$\partial \dddot{x}/_{\partial x} = \partial \dddot{x}/_{\partial \dot{x}} = \partial \dddot{x}/_{\partial \ddot{x}} = f_{31} = f_{32} = f_{33} = 0 \tag{184}$$

$$F = \begin{bmatrix} 0 & -1 & 0 \\ f_{21} & f_{22} & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{185}$$

$$\Phi_{1\,term} = I + FT \tag{186}$$

$$\Phi_{2\,term} = I + FT + \frac{1}{2}F^2T^2 \tag{187}$$

$$Q = \Phi_s \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}; Q_k = \int_0^{Ts} \Phi(\tau) Q \Phi(\tau)^T d\tau \tag{188}$$

$$\hat{x}(3) = \frac{-\gamma e^{-\eta \hat{x}(1)} g (\hat{x}(2))^2}{2\beta} + g \tag{189}$$

The Φ and Q_k matrices become verbose, and are calculated using the symbolic math toolbox in the Matlab. During the prediction step equation (188) is used to supplement fundamental matrix with the more accurate acceleration estimate.

5.3.1.4 Particle Filter with Non-Linear Acceleration

There are no matrices to derive when implementing the particle filter. Equations (176) and (199) describe how system states were propagated for the prediction step of the filter. By itself, initializing the noise terms in a particle filter can be challenging. Approaches for remedying this usually involve algorithms applied outside of the particle filtering framework, such as an optimization [Bréhard 2004, Wang 2007]. To make a fair comparison a simple implementation of the particle filter known as sequential importance resampling (SIR) is used for benchmarking, and a guess and check strategy was used to determine reasonable noise terms for initialization.

5.3.2 Comparison of Tracking Results for a Known Ballistic Object

The case of tracking a known ballistic object is a very simplistic scenario and not very realistic. It is unlikely in practice that you would have a perfect model that incorporates all inputs that affect the dynamics of a system. Furthermore it is unreasonable to think that you would have a priori knowledge of the free parameters in the model (drag coefficient β) and the initial state of the system. None the less it provides a foundation for comparing the performance of the different algorithms implemented in this work.

Figure 53 compares the results of applying the ballistic object tracking benchmark data set using various recursive algorithms. For each algorithm four measurement frequencies were simulated (0.1, 0.01, 0.005, and 0.001 seconds/measurement) to investigate the effect that linearization in the EKF has on performance. For the Kalman family of filters implementation with both a one term and two term approximation of the fundamental matrix is investigated. Each algorithm was initialized with the same noise terms, $R = (7.62m)^2, \Phi_s = (10g)^2/30$, and the same initial conditions, $\hat{x}_0 = \lfloor 61000\,m \quad 3048\,m/sec \quad 10\,\mathrm{m/sec^2} \rfloor^T$ which is the same as

the simulation. The PF requires additional noise terms to avoid particle degeneracy. The noise terms for the position, velocity and acceleration in the PF were set as $(7.62m)^2$, 10 m/sec, and 10 m/sec^2. The standard deviation for the penalty function in the PF was set at 10 m. The particle population size for the PF was 30. As derived above, the Kalman filter was implemented with a simple position, velocity, acceleration state vector. The EKF and PF were implemented with position, velocity and non-linear acceleration. Because this is a benchmark test on a simulated data set it is possible to evaluate performance for the observed position estimate of the ballistic object as well as the unobserved velocity and acceleration estimates. For the Kalman family of filters the root mean squared error between the estimated state and the actual (simulated) state are used as a measure of performance. In theory if a Kalman filter is operating correctly, about 63% of the residuals between the estimated and true value should fall inside the uncertainty bounds prescribed in the covariance matrix. To determine the theory bounds for the filter the square root of the position term in the covariance matrix, $P_k(1,1)$,is calculated. If the absolute value of the residual between the estimated position and the true (simulated) position is less than $\sqrt{P_k(1,1)}$, then the estimate is considered inside the theoretical error bounds. This calculation is repeated for every measurement step, and reported as a percentage of estimates that fell inside the theory bounds. This calculation can also be performed using the residual between the projected state and measured state of the filter. For the Kalman filter, this is one of the few known methods for monitoring for filter divergence online.

				RMS of Error			
Algorithm	State Vector	Terms	Time Step	Position [m]	Velocity [m/s]	Acceleration [m/s^2]	Percent Inside [%]
KF	PVA	1	0.1	4.81	12.37	22.26	60.80
KF	PVA	1	0.01	1.70	6.39	16.60	68.58
KF	PVA	1	0.005	1.43	5.45	15.49	64.94
KF	PVA	1	0.001	0.72	3.49	13.68	66.94
KF	PVA	2	0.1	4.81	12.37	22.26	60.80
KF	PVA	2	0.01	1.70	6.39	16.60	68.58
KF	PVA	2	0.005	1.43	5.45	15.49	64.94
KF	PVA	2	0.001	0.72	3.49	13.68	66.94
EKF	PVNA	1	0.1	3.75	6.51	11.60	76.74
EKF	PVNA	1	0.01	1.59	3.97	10.04	71.74
EKF	PVNA	1	0.005	1.23	3.42	9.93	71.32
EKF	PVNA	1	0.001	0.66	2.28	9.84	71.50
EKF	PVNA	2	0.1	3.73	6.21	11.58	77.08
EKF	PVNA	2	0.01	1.59	3.94	10.04	71.81
EKF	PVNA	2	0.005	1.23	3.41	9.93	71.35
EKF	PVNA	2	0.001	0.66	2.28	9.84	71.52
PF	PVNA	N/A	0.1	16.92	56.45	16.06	N/A
PF	PVNA	N/A	0.01	8.35	141.11	23.11	N/A
PF	PVNA	N/A	0.005	8.34	199.79	40.25	N/A
PF	PVNA	N/A	0.001	8.28	561.68	62.28	N/A

Eror is estimated state minus actual state (no noise)
KF = Kalman Filter
EKF = Extended Kalman Filter
PF = Particle Filter
PVA = Position, Velocity, Acceleration
PVNA = Position, Velocity, Non-linear Acceleration
Terms = Terms in the approximation of the fundamental matrix

Figure 53: Benchmark Testing for Various Implementations of recursive filters

A few points to notice in the comparison of tracking a known ballistic object

- Results for the one term and two term Kalman filters are negligibly different
- The percent improvement in the EKF improves as the sampling time decreases
- The residuals of the EKF are inside the theory bounds more often than the KF
- The PF is more dependent on sampling time than the EKF

- The choice of noise terms for the PF is a function of the sampling time. In this comparison the noise terms were optimized for the 0.1 seconds/measurement simulation and never modified for other sampling rates.

This concludes the comparison of the KF and EKF for the very simple case of a constant known ballistic coefficient. In more realistic situation the ballistic coefficient is not a known value. The next section will address estimating the ballistic coefficient, which is itself an intermediate step towards tracking an object with a time varying ballistic coefficient.

5.3.3 Tracking an Unknown Object (β constant, with unknown a priori value)

The more realistic case of tracking an object with an unknown ballistic coefficient is demonstrated using the PVA state vector for the KF, and a position, velocity, ballistic coefficient (β-parameter) state vector with the EKF and PF. This state vector is abbreviated as the PVB state vector.

$$x_k = \begin{Bmatrix} x \\ \dot{x} \\ \beta \end{Bmatrix} \tag{190}$$

Similar to previous cases the ballistic coefficient is assumed to be a constant, and the derivative of the ballistic coefficient is taken to be white noise with a standard deviation of u_s. The differential equations become

$$\begin{aligned} \dot{x} &= \quad -\dot{x} \\ \ddot{x} &= -(D(x_k) - g) = \frac{-\gamma e^{-\eta x} g \dot{x}^2}{2\beta} + g \\ \dot{\beta} &= \quad u_s \sim N(0, \Phi_s) \end{aligned} \tag{191}$$

The units and uncertainty surrounding the beta term are different than the Φ_s used previously for the non-linear acceleration state. If the ballistic coefficient can be assumed to be chosen to within +/- 1000 kg/m-s^2, then the process noise term can be set at $\Phi_s = (10000)^2/30 = 33333$. The terms for the Jacobian become

$$\partial\dot{x}/_{\partial x} = f_{11} \tag{192}$$

$$\partial\dot{x}/_{\partial\dot{x}} = f_{12} = -1 \tag{193}$$

$$\partial\dot{x}/_{\partial\beta} = f_{13} = 0 \tag{194}$$

$$\partial\ddot{x}/_{\partial x} = f_{21} = \frac{\eta\gamma e^{-\eta x} g\dot{x}^2}{2\beta} \tag{195}$$

$$\partial\ddot{x}/_{\partial\dot{x}} = f_{22} = \frac{-\gamma e^{-\eta x} g\dot{x}}{\beta} \tag{196}$$

$$\partial\ddot{x}/_{\partial\beta} = f_{23} = \frac{\gamma e^{-\eta x} g\dot{x}^2}{2\beta^2} \tag{197}$$

$$\partial\dddot{x}/_{\partial x} = \partial\dddot{x}/_{\partial\dot{x}} = \partial\dddot{x}/_{\partial\beta} = f_{31} = f_{32} = f_{33} = 0 \tag{198}$$

$$F = \begin{bmatrix} 0 & -1 & 0 \\ f_{21} & f_{22} & f_{23} \\ 0 & 0 & 0 \end{bmatrix} \tag{199}$$

$$\Phi_{1\,term} = I + FT \tag{200}$$

$$\Phi_{2\,term} = I + FT + \frac{1}{2}F^2T^2 \tag{201}$$

$$Q = \Phi_s \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix};\ Q_k = \int_0^{Ts} \Phi(\tau) Q \Phi(\tau)^T d\tau \tag{202}$$

5.3.4 Comparison of Tracking Results for a Unknown Object

The case where the characteristics of the ballistic object are not known a priori makes the simulation more realistic. The beta coefficient is a model term that is desired to be estimated online in a joint state parameter estimation framework. To realize joint state parameter estimation the beta coefficient has been integrated as one of the system states being estimated by the filter. It is assumed that the beta coefficient is a constant. Figure 54 and Figure 55 show the tracking results for the EKF and PF implementations of the PVB filter. In both filters the beta parameter was incorrectly initialized as $\beta = 10000\ kg/(m - sec^2)$. In the simulation $\beta =$

19161 $kg/(m-sec^2)$. The EKF's estimate of the beta parameter (xk[3]) increased during the simulation, but never converged on the true value. The PF implementation did not show any trend towards estimating the correct parameter. The failure of the PF to jointly estimate states and parameters is not surprising for the simple PF implementation. Modifications to the PF for performing joint state and parameter estimation are discussed in [Saha 2011]. A chart comparing each of the different filters is found in Figure 56.

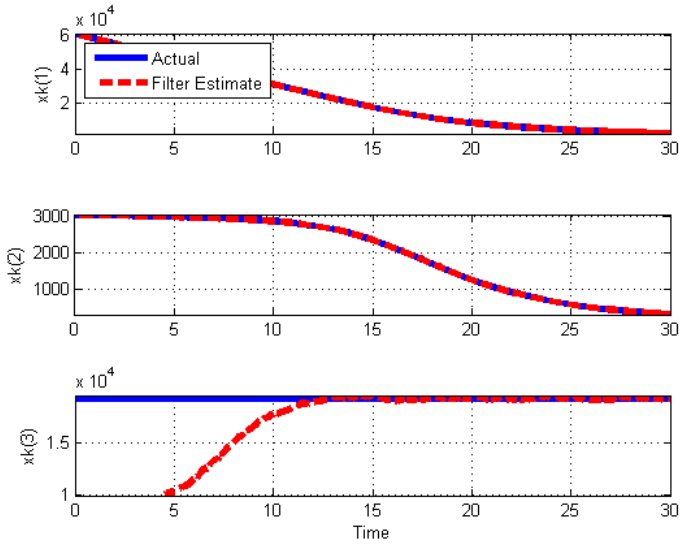


Figure 54: Tracking results for EKF PVB filter with a poorly initialized ballistic coefficient

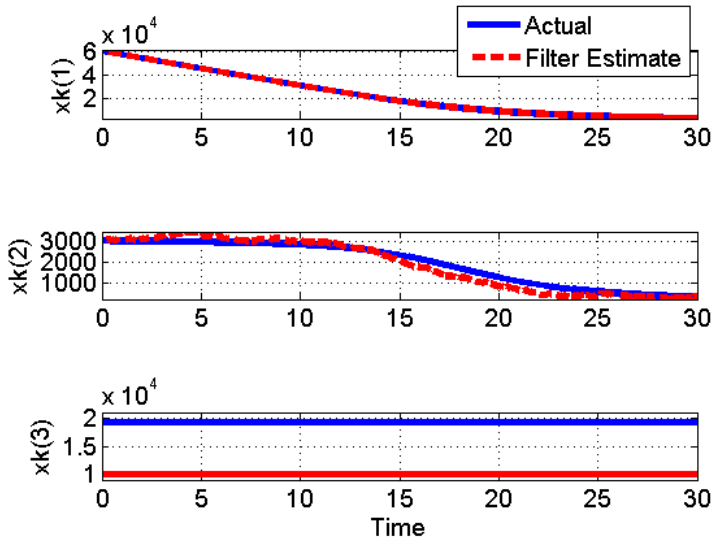


Figure 55: Tracking results for PF PVB filter with a poorly initialized ballistic coefficient

				RMS of Error				
Algorithm	State Vector	Beta Initialized [kg/(m-sec^2)]	Time Step [sec/meas.]	Position [m]	Velocity [m/s]	Acceleration [m/s^2]	Beta [kg/(m-sec^2)]	Percent Inside [%]
KF	PVA	19161	0.01	1.70	6.39	16.60	N/A	68.58
EKF	PVB	19161	0.01	0.80	1.71	N/A	15.96	71.91
PF	PVB	19161	0.01	8.42	231.49	N/A	32.64	N/A
KF*	PVA	10000	0.01	1.70	6.39	16.60	N/A	68.58
EKF	PVB	10000	0.01	2.47	3.45	N/A	4377.10	67.54
PF	PVB	10000	0.01	8.39	270.93	N/A	9177.10	N/A

Eror is estimated state minus actual state (no noise)
KF = Kalman Filter
EKF = Extended Kalman Filter
PF = Particle Filter
PVA = Position, Velocity, Acceleration
PVB = Position, Velocity, Ballistic Coeficient (Beta)
Beta Actual = 19161 kg/(m-s^2)
*Result is independent of choise of beta, exactly same result as first KF implementation

Figure 56: Tracking performance for PVB filters with the beta parameter initialized correctly and incorrectly

A few points to notice in the comparison of tracking an unknown ballistic object

- Performance of the EKF PVA filter (Figure 53) is inferior to the PVB filter for a well initialized filter.

- The KF implementation is independent from the choice of β and performs better at tracking in the case where the ballistic coefficient of the object being tracked is not well known.
- For the case of a poorly initialized filter, the KF slightly out performs the EKF
- The EKF can perform joint state parameter estimation without modifications to the state vector or filtering framework
- Note how in Figure 57 the estimate of the ballistic coefficient does not stabilize until the impact of air resistance becomes numerically significant.

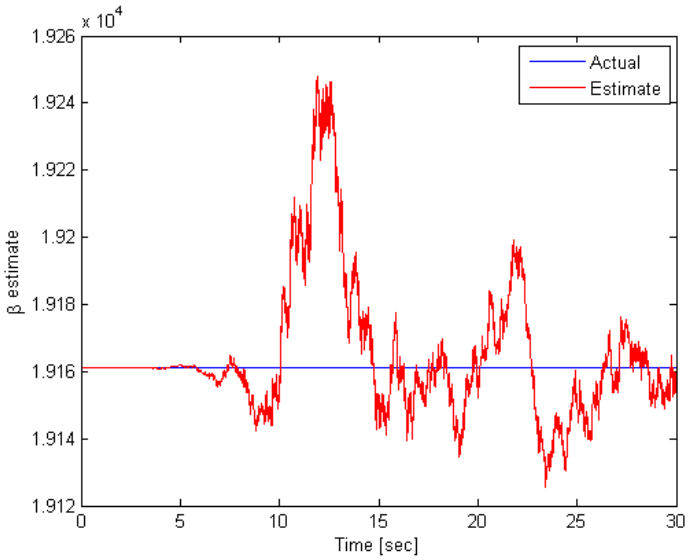


Figure 57: Estimate of ballistic coefficient. Note the very fine scale on the y-axis.

If the only criteria for judging the performance of a filter is the quality of the tracking, the PVA and PVB perform nearly identically in this example for well initialized filters. In practice with prognostic applications the ability to perform joint state and parameter estimation can have a significant effect on the quality of a prognostic algorithms performance.

The KF has a very low error when tracking, but is simply data driven, the ballistic coefficient is not part of the update equations. An estimate of the ballistic coefficient can be made by re-arranging the governing equation for acceleration.

$$\beta = \frac{g\rho\dot{x}^2}{2(\ddot{x} + g)} \tag{203}$$

Unfortunately since the estimated acceleration is not very smooth the estimate of the ballistic coefficient is not very smooth and off by a couple orders of magnitude in some locations.

The previous discussion has established a methodology for quantifying the operation of a filter, and shown how uncertainty in initial parameters effect the tracking and parameter estimation of an algorithm. The final section will discuss the situation when the model parameters are time evolving.

5.3.5 Tracking an object with time evolving parameters

In this section the ballistic coefficient is initially unknown, and additionally is time evolving. The simulation has been changed to represent the case where at 15 seconds the object being tracked splits in half causing the ballistic coefficient to decrease by 50%. The PVA KF, and PVB EKF and PF are used for tracking

The effect of the change in the ballistic coefficient on the true state vectors is shown below.

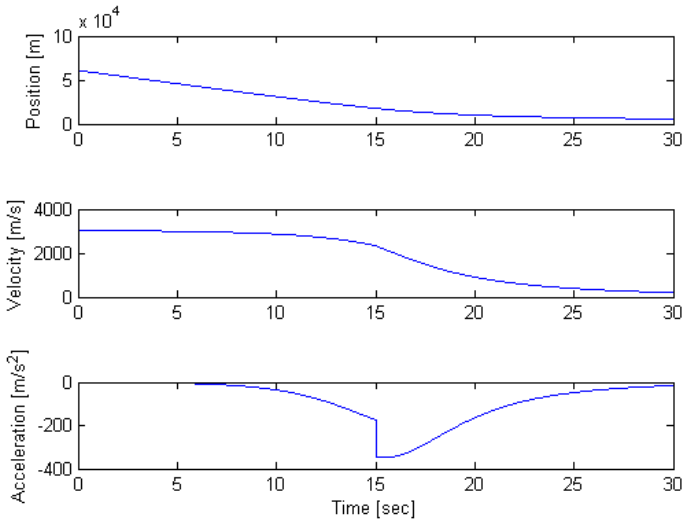


Figure 58: Simulation with the added complexity of a time varying drag coefficient

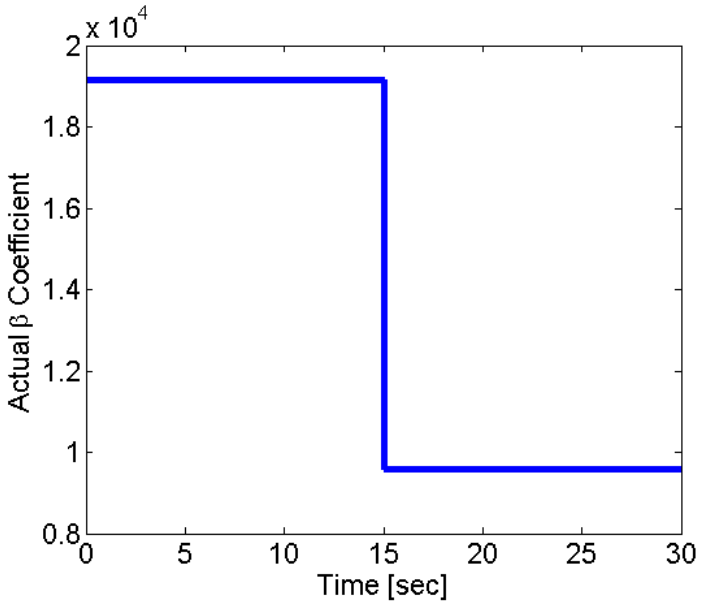


Figure 59: Change in the simulated beta coefficient. The value is reduced by 50% for the final 15 seconds of the simulation.

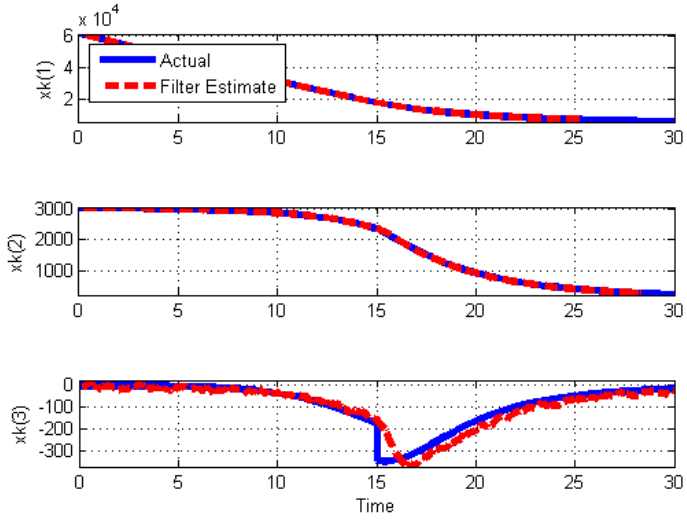


Figure 60: Tracking performance for the KF with an evolving beta parameter

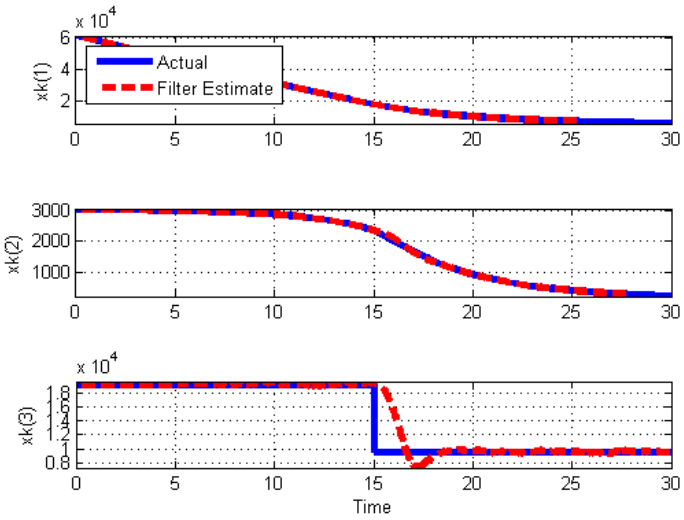


Figure 61: Tracking performance for the EKF with an evolving beta parameter

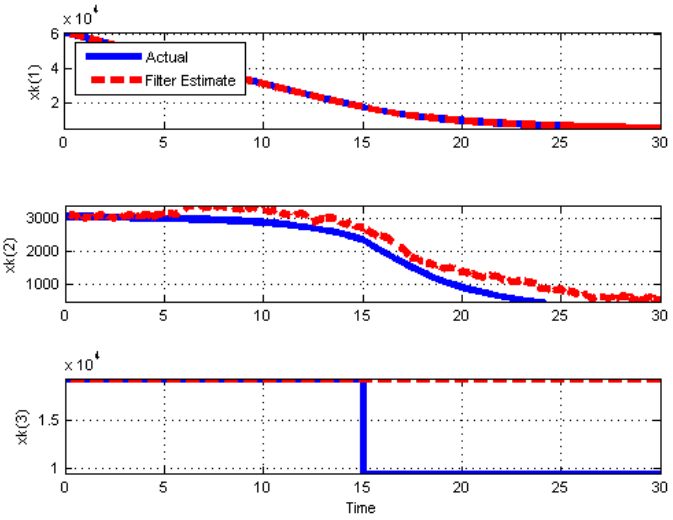


Figure 62: Tracking performance for the PF with an evolving beta parameter

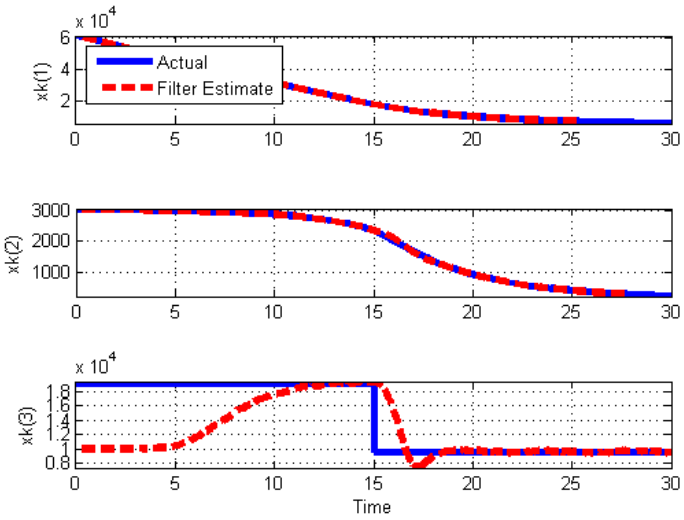


Figure 63: Tracking performance for a poorly initialized EKF with an evolving beta parameter

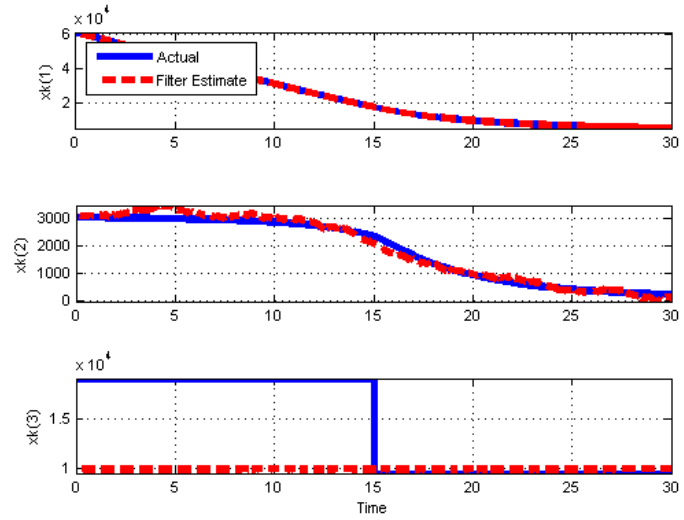


Figure 64: Tracking performance for a poorly initialized PF with an evolving beta parameter

				RMS of Error				
Algorithm	State Vector	Beta Initialized [kg/(m-sec^2)	Time Step [sec/meas.]	Position [m]	Velocity [m/s]	Acceleration [m/s^2]	Beta [kg/(m-sec^2)]	Percent Inside [%]
KF	PVA	19161	0.01	2.24	10.85	32.13	-	58.48
EKF	PVB	19161	0.01	6.20	19.97	-	1677.20	62.05
PF	PVB	19161	0.01	8.47	360.93	-	6759.90	-
EKF	PVB	15000	0.01	6.59	20.13	-	4685.30	35.69
PF	PVB	15000	0.01	8.40	172.69	-	6494.50	-

Eror is estimated state minus actual state (no noise)
KF = Kalman Filter
EKF = Extended Kalman Filter
PF = Particle Filter
PVA = Position, Velocity, Acceleration
PVB = Position, Velocity, Ballistic Coeficient (Beta)
Beta Actual = 19161 kg/(m-s^2) for 15 seconds, 9581 for 15 seconds
*Result is independent of choise of beta, exactly same result as first KF implementation

Figure 65: Comparison of recursive filters for a time evolving beta parameter

By simply looking at the tracking accuracy the PVB filter appears to be the worst performer, but it will be shown that it is actually the best filter for prognosticating RUL. The advantage of the EKF-PVB filter is the ability to perform joint state and parameter estimation which can be seen by looking at the comparison of the actual ballistic coefficient and estimated coefficient.

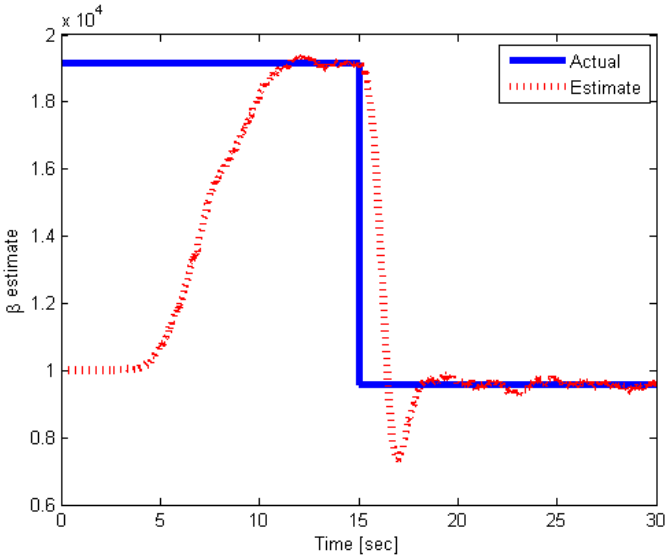


Figure 66: Filter estimate of the drag coefficient

5.4 Conclusion

This chapter presented a number of methods for estimating the state of a system. To provide context, batch least squares methods were introduced, then a recursive least squares approach. It was shown that the recursive estimate matched with the batch estimate at the very last data set. Savitzky-Golay smoothing was discussed as a possible least squares based algorithm for prognostics. Three state estimation techniques, the Kalman filter, Extended Kalman filter, and particle filter were introduced. A large number of simulations were conducted to highlight the performance of different filters when subjected to different scenarios. Specifically the effect of approximations in the fundamental matrix, measurement sampling time, initialization states, and joint state parameter estimation were conducted. Each filter has its own strengths and weaknesses which were highlighted. Recursive filtering is only part of a PHM algorithm, and the benefits of applying each algorithm to real data sets, as opposed to the simulations presented in this section will be performed in the following sections.

6 Quantifying performance of PHM implementations

Methods for quantifying the performance of PHM implementations are discussed in this section. A review of traditional statistical time series metrics and their shortcomings when applied to prognostics are discussed in [Saxena 2008a,b, 2009a,b]. Definitions and metrics used to describe the performance of a PHM implementation are introduced through an example PHM implementation on a simulated data set. The same set of metrics will be used to quantify performance of various PHM implementations implemented on actual systems in the proceeding sections and provides a standard measure of performance across widely disparate setups and experimental conditions.

6.1 Example PHM Implementation

Hypothetically assume you can monitor a feature of a system that is incurring physical damage from an environmental stress, and believe the damage degradation to follow a quadratic model.

$$y = x^2 \tag{204}$$

Define failure as the time when the feature vector y is at or above a value of 4. A measurement of the system is represented as the true system y, with additive white noise of standard deviation σ.

$$z = y + N(0, \sigma) \tag{205}$$

A time history (created in simulation) of the measured system as it progresses from a health state at level zero, to a failed state at level 4 is shown in Figure 67.

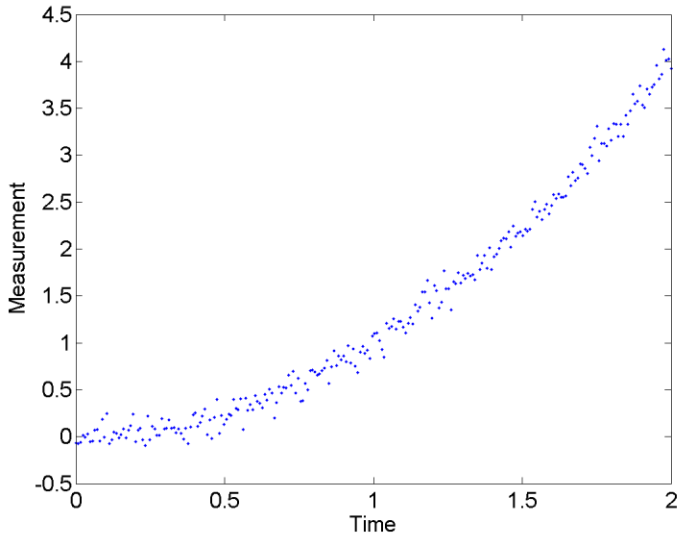


Figure 67: Noisy measurement. Units are purposefully omitted to avoid accidently mistaking this data as a real experimental observation.

The goal is to predict the expected time of failure, or the point when the feature will have a value greater than four. The expected time of failure is also referred to as the remaining useful life of the system. Estimates of the remaining useful life to will be updated after every 25 measurements until failure occurs. A Kalman filter is used to recursively process each new measurement as it becomes available. The Kalman filter was initialized as a three state filter. The feature vector is the measured state. The first and second derivative of the measured state are also used as state variables.

$$\hat{x} = \begin{Bmatrix} y \\ \dot{y} \\ \ddot{y} \end{Bmatrix} \tag{206}$$

In state space form the system is represented as

$$\begin{Bmatrix} \dot{y} \\ \ddot{y} \\ \dddot{y} \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} y \\ \dot{y} \\ \ddot{y} \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ u_w \end{Bmatrix} \tag{207}$$

It is assumed that $\ddot{y}$ is a constant and therefore its derivative should be white noise. A two term approximation is utilized for the fundamental matrix.

$$\Phi_{2\,term} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} t + \frac{1}{2}\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} t^2 \tag{208}$$
$$= \begin{bmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}$$

The simulation was created by adding white noise with a standard deviation of 0.1 to the true underlying system. The Kalman filter will be initialized with a noise term R that has a standard deviation of 0.1. The process noise variance w is estimated by the assumption that the feature vector will at most change from zero to four over a time period of two units of time.

$$u_w = \frac{4^2}{2} \tag{209}$$

The time step for the simulation and filter will be 0.008 units of time. Initial state values were initialized as zero due to ignorance of how the system might actually be operating. To make a prediction of the remaining useful life the current state estimate provided by the Kalman filter will be combined with the expected future behavior of the system stated in equation (204). The failure threshold x_f is defined as four. Using the current estimated state of the system, and the failure threshold a quadratic equation can be written.

$$x_f = \hat{x}_1 + \hat{x}_2 t_{RUL} + 0.5\hat{x}_3 t_{RUL}^2 \tag{210}$$

Solving the proceeding equation for t_{RUL} provides an estimate of the remaining useful life at the time of the prediction. As new measurements become available before the failure of the system, the same equation can be utilized with the most recent estimate of $\hat{x}$ to calculate updated remaining useful life. The uncertainty (standard deviation) surrounding the prediction is calculated using diagonal terms from the posterior error covariance matrix P [Swanson 2001].

$$\sigma_{RUL} = \frac{P(1,1)}{P(2,2)} \tag{211}$$

For completeness the tracking and filtering results from the Kalman filter are shown in

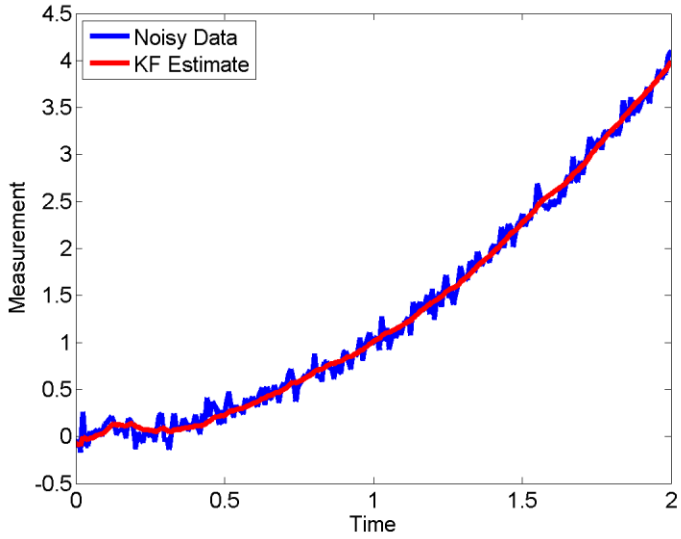


Figure 68: Tracking results from the Kalman filter for the synthetic data set

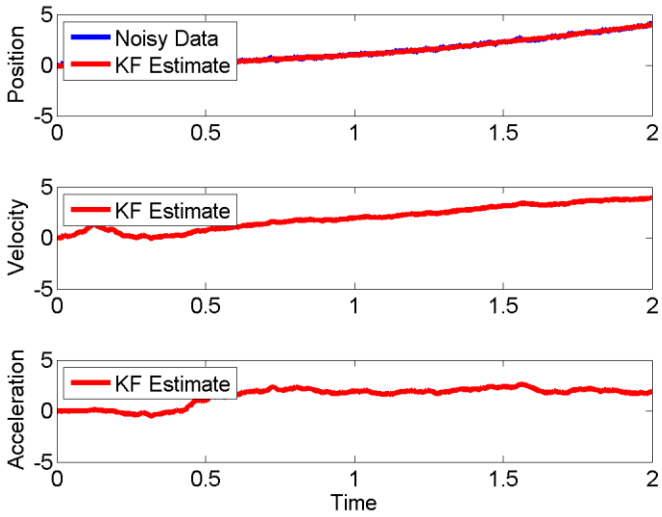


Figure 69: The Kalman filter successfully tracked the noisy data and provides smooth estimates of the first two derivatives. As initially assumed, the second derivative is approximately constant

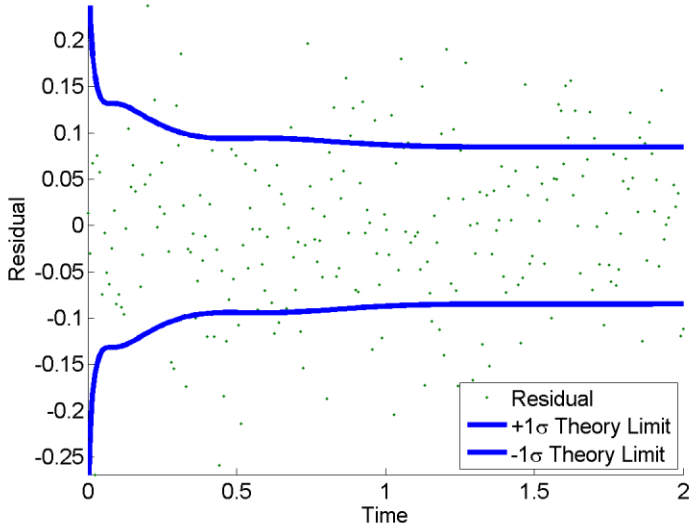


Figure 70: Residuals between the estimated and actual state estimates falls within the theoretical error bounds 66% of the time and indicates that the Kalman filter is operating correctly. Unfortunately this provides no insight if RUL predictions are accurate and correct

6.2 Prognostics Hindecaste

This section will present a proof of concept for a prognostic system to monitor and manage the health of a fictitious system that was simulated in the previous section. By simulating the data set the true underlying state that generated the noisy data is known, and makes performance evaluation easier. In practice the true state is not possible to obtain. To borrow terminology from hurricane forecasting, a forecast is the real time prediction of future events based on observations. A hindecaste is a simulated real time prediction of future events, using previously observed data. New hurricane models are validated using the hindecaste method on more than one hundred years of climatology data [Rappaport 2009]. A major pitfall of this approach is inadvertently creating a future information leak. Future information leak is the use of information for a prediction that is not yet available. An example of a future information leak is a three day moving average centered on the current day. When designing and testing PHM algorithms is preferable to simulate performance so you do not have the time delay or expense of

running a system to failure. Unfortunately this method can also result in a implementation that only performs good on the available data set, but does not perform well on real data. Methods for designing more robust PHM algorithms is discussed in the particle swarm optimization section. The simulated data set used in this example simply provides an easy discussion of the relevant ways to quantify the performance of a PHM algorithm.

Illustrating the prediction process is difficult because a prediction without a frame of reference cannot be judged as successful. Therefore when discussing the example problem we commit the logical fallacy of a future information leak on purpose. Figure 71 shows an illustration of an early prediction from the PHM algorithm described earlier. The upper plot shows the data that was available at the time of the prediction overlaid on clean system state. In practice the clean system state is not available. To provide a frame of reference for the prediction the lower plot is an overlay of the prediction on top of the full simulated data set. Again this would not be possible to create if this was a true forecast, and not a simulated hindecaste.

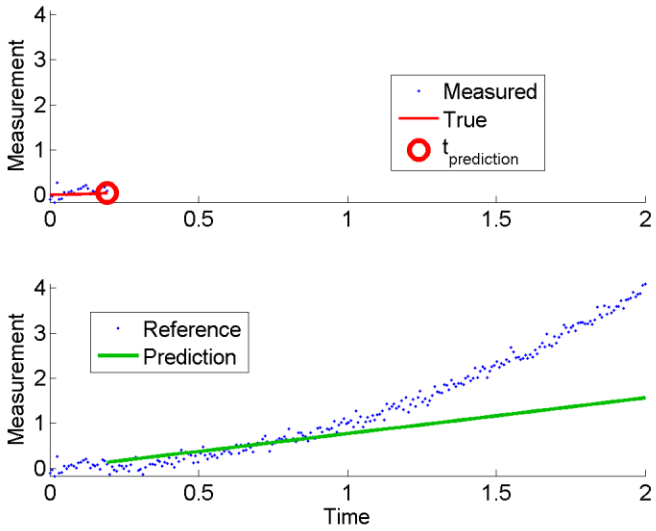


Figure 71: Illustration of the Prediction Process with an Overlay on the reference data set. With little trend in the measured data provides little insight into the nature in which damage will propagate in the system

After every 25 measurements the prediction process described in equation (210) is repeated with the most recent state estimate from the Kalman filter. The next eight predictions are shown in a similar manner.

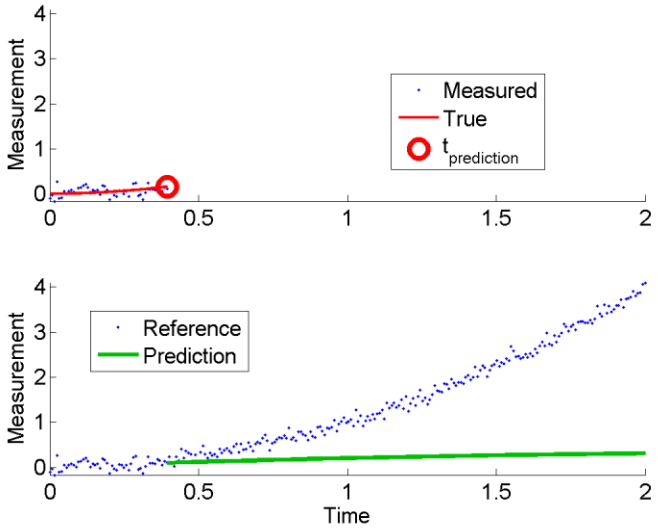


Figure 72: Prediction after 50 measurements. With an observation of available data that is not above the noise floor, the prediction of RUL is significantly over estimated

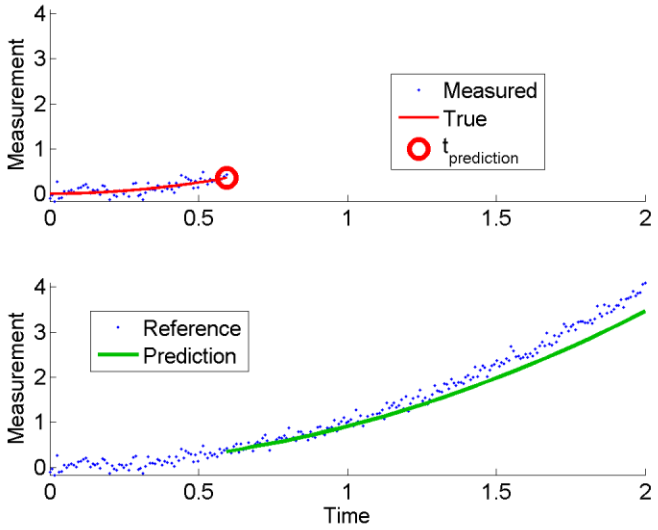


Figure 73: Prediction after 75 measurements. At this point in time enough measurements have been made to more accurately estimate the form of the underlying process that is generating the noisy data. The algorithm is still slightly overestimating the actual RUL

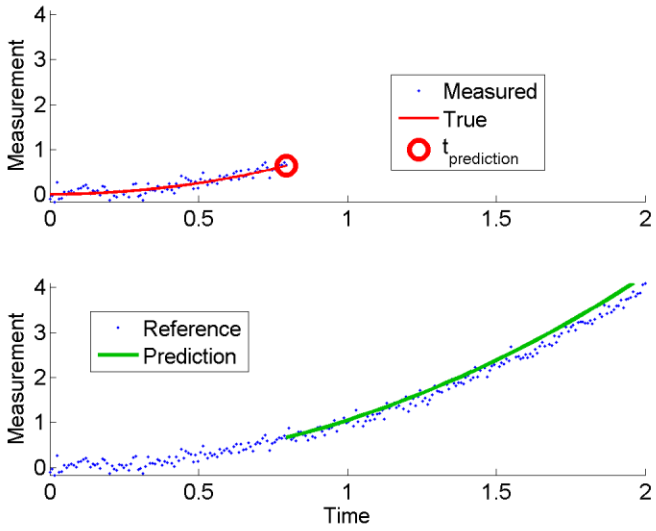


Figure 74: Prediction after 100 measurements. In contrast to the previous predictions, at this point in time the RUL is slightly under predicted.

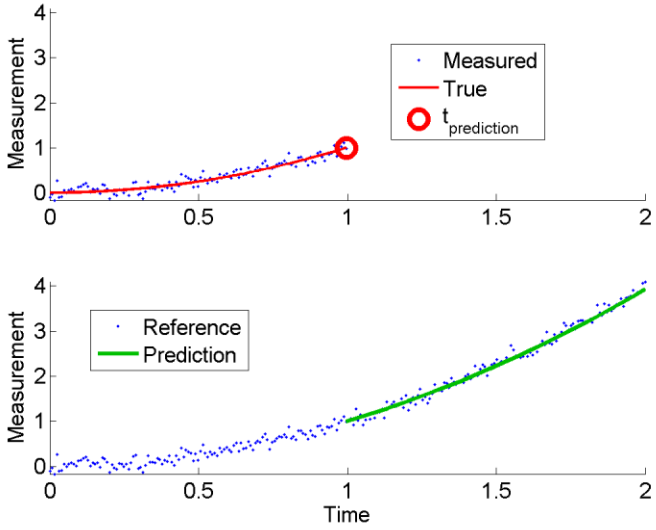


Figure 75: Prediction after 125 measurements. This prediction represents the first prediction that has a very high accuracy compared to the actual RUL

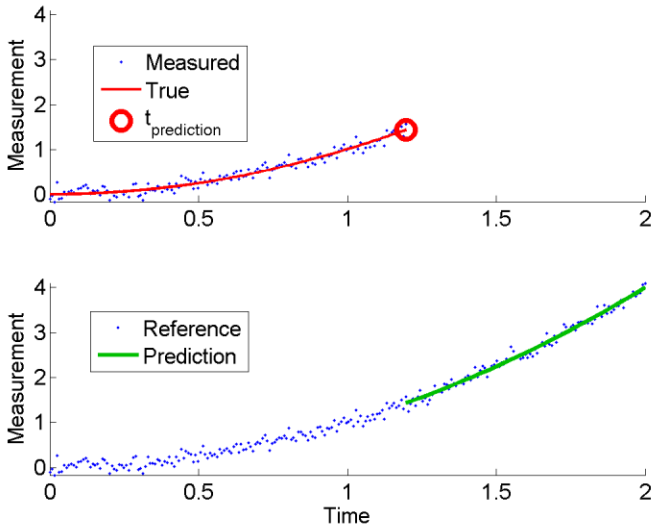


Figure 76: Prediction after 150 measurements. At this point in the simulation it is apparent that the PHM algorithm has converged to the actual RUL.

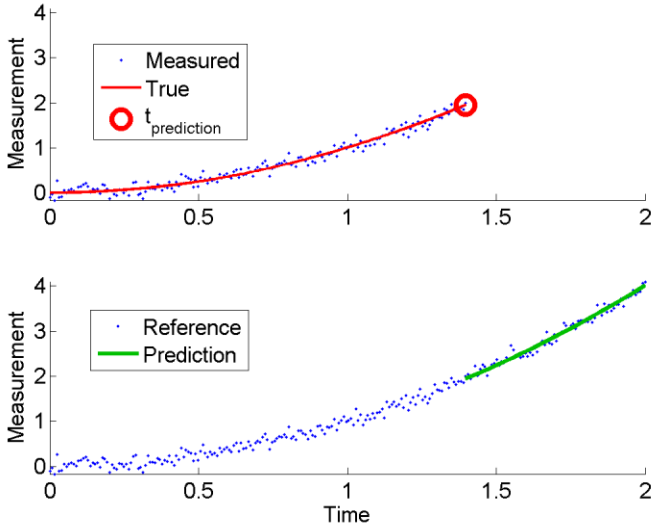


Figure 77: Prediction after 175 measurements. Even though the PHM algorithm has converged to a accurate prediction updates are still made in the event that the underlying process is affected by an un-modeled input

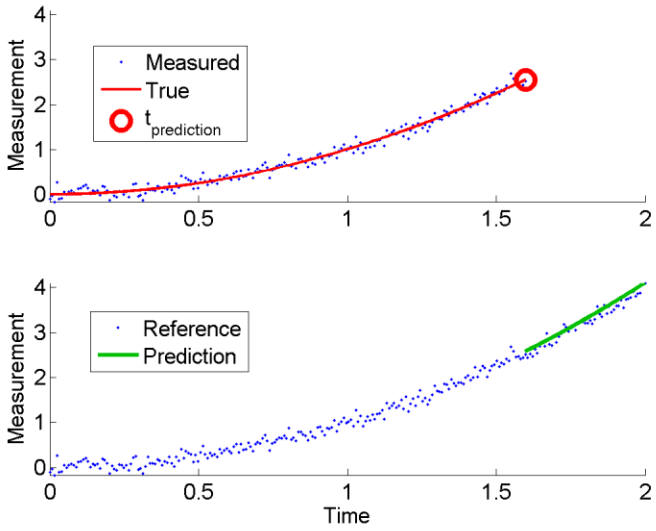


Figure 78: Prediction after 200 measurements. During the final stages of the simulation the fictitious product is incurring damage at a much faster rate.

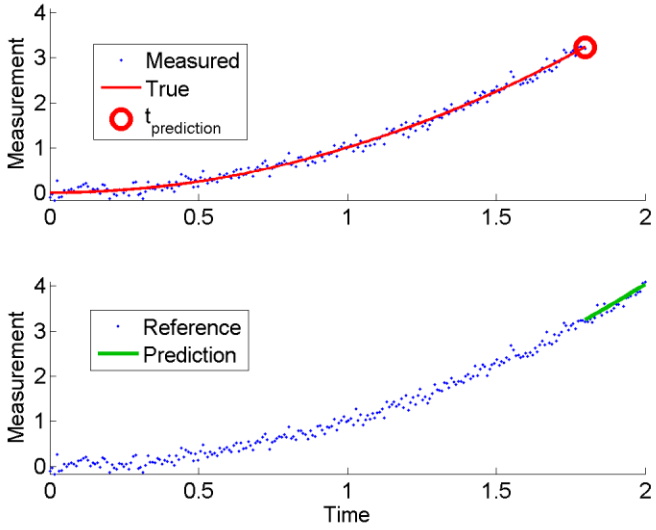


Figure 79: Prediction after 225 measurements. The final prediction before failure occurs. For such short time scales the specific shape of the failure propagation curve is no longer important. A linear model would perform nearly as well as a quadratic model.

Reporting prediction results in the manner shown above are impossible for implementation that are making predictions in real time. Even if you are using previously recorded data in a hindecaste the density of information is very low, and not very easily converted into a quantitative analysis. To summarize the accuracy of the predictions presented

above a remaining useful life plot can be created. The horizontal axis of the plot is time. The vertical axis of the plot is remaining useful life. Only after a system has been run to failure is the actual remaining useful life observable. On the RUL plot the actual remaining useful life is a straight line. The predicted RUL is also plotted on the same scales to compare the accuracy of the predictions to the actual RUL. It is implicitly assumed that the stresses that caused the system to fail were relatively constant or repeatable. There are a number of cases where this type of comparison would be less useful, but currently there are no known alternatives and the metric has been historically used to communicate results from PHM algorithms. Figure 80 shows a very simple implementation of an RUL plot as an illustrative example. Figure 81 summarizes the predictions shown in Figure 72 through Figure 79 along with predictions made after every new data point that were not shown. The light pink band is a user defined band of acceptable prediction accuracy that will facilitate quantifying performance. The alpha bounds are simply $\pm\alpha$ percent of the actual remaining useful life and are considered arbitrary based on the required performance of the PHM implementation. Because the actual remaining life is a number that decreases with time, the alpha bounds consequently become tighter as the point of failure draws closer.

Real Time		After Failure
Elapsed Time [Hr]	Predicted RUL [Hr]	Actual RUL [Hr]
0	8	7
1	6	6
2	4	5
3	3	4
4	2	3
5	2	2
6	2	1
7	1	0

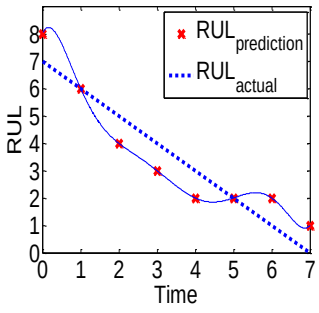


Figure 80: An illustrative example to demonstrate how RUL plots are calculated

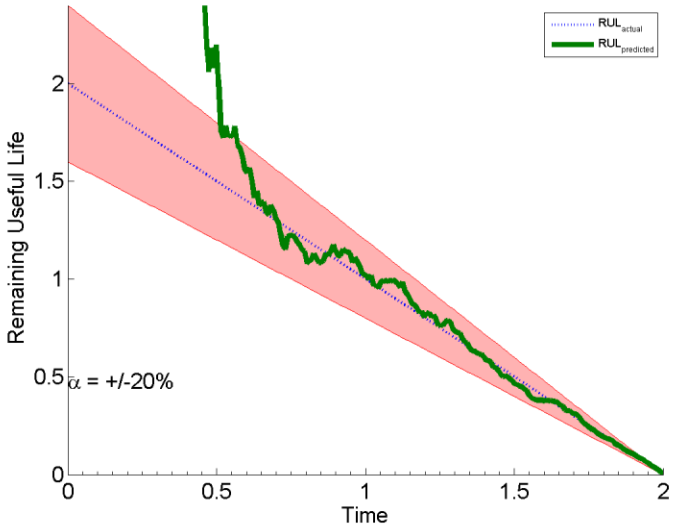


Figure 81: Summary of RUL predictions for the hindcaste on the fictitious data set

6.3 PHM Specific Offline Performance Metrics

The alpha lambda plot is the starting point for a family of offline PHM performance metrics introduced in [Saxena 2008a,b, 2009a,b]. The RUL plot successfully illustrates the accuracy of the PHM predictions, but is more difficult to use to interpret the precision of the predictions. To provide more insight into algorithm performance a metric known as the β metric is introduced to quantify precision. The relative accuracy metric is considered a more meaningful description of accuracy than simply the error between the predicted and actual RUL. Relative accuracy penalize errors closer to the end of life more harshly than errors at the beginning of the prediction time. Each of the metrics are described in more detail in this section. The metrics are considered offline metrics because they require the system to fail before being calculated. Unfortunately this metrics can provide no real time insight into the accuracy of the RUL predictions.

6.3.1 Alpha-Lambda Metric

In addition the predicting remaining useful life, risk based decision making schemes require an estimate of the uncertainty surrounding the prediction. This is reported as the standard deviation for the RUL prediction in equation (211). The uncertainty information can be communicated on an alpha-lambda plot. The alpha-lambda plot is closely related to the RUL plot, but the horizontal axis is replaced with the normalized time to failure and the uncertainty bounds for each prediction are also plotted. A normalized time scale is used to facilitate comparisons between different tests that may have different run to failure times.

$$\lambda = t/t_{end} \tag{212}$$

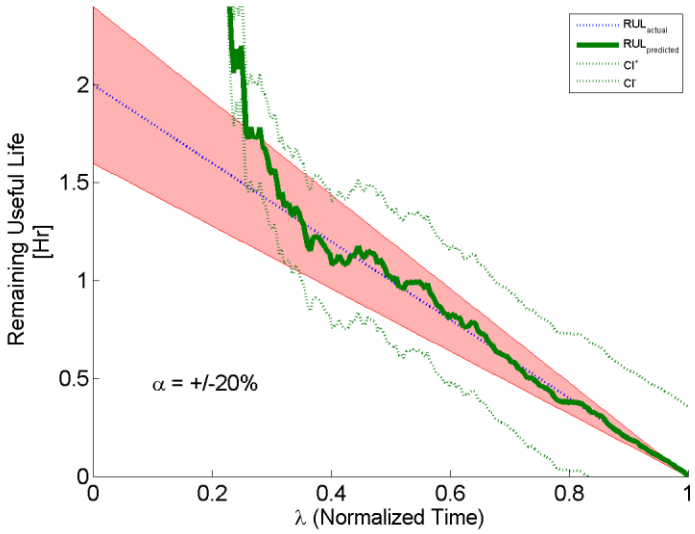


Figure 82: Alpha-Lambda plot which shows both RUL predictions and the uncertainty for each prediction

6.3.2 β Metric

Used to quantify the precision of a predicted RUL compared to the actual RUL, the β metric is calculated differently depending on the method used to describe uncertainty. In

algorithms based on the Kalman filter or the extended Kalman filter it is implicitly implied that all values are Gaussian in nature.

6.3.2.1 Gaussian β metric

In the case of a normal distribution only the mean μ and standard deviation σ are needed to fully describe the distribution. In this work the mean is the estimate of remaining useful life and the standard deviation is calculated from equation (211) or a similar method. In a normal distribution the area under the distribution is defined as unity. The β metric is evaluated as the percentage of the area under the normal distribution that falls within the user defined alpha bounds. Predictions where most of the probability mass represented by the distribution is inside the alpha bounds represent a prediction that is successfully achieving the precision required by the user. A prediction that is accurate, but not very precise is not very helpful when making decisions in the risk based framework. If only a small amount, or none, of the probability mass is inside the alpha bounds the prediction is either not accurate or not precise. The probability mass inside the alpha bounds is represented as

$$\beta_k = \int_{-\alpha}^{+\alpha} PDF(x)dx \tag{213}$$

The PDF, in this case a normal distribution, is integrated between the alpha bounds to evaluate the metric. At each prediction time the PDF may be different and the β metric may have a different value. Due to the decreasing size of the alpha bounds closer to the point of failure the beta metric score will fall off even if the uncertainty has converged to a constant value.

To calculate the β metric for a normal distribution a cumulative distribution function is used. The area under a normal distribution between negative infinity and an arbitrary value a is defined as

$$D(a) = 0.5\left(1 + erf\left(\frac{a-\mu}{\sigma\sqrt{2}}\right)\right) \tag{214}$$

Where erf is the error function , μ is the mean or predicted RUL, and σ is the standard deviation of uncertainty in the predicted RUL. To calculate the area between two points, such as the upper alpha bound $\alpha^{(+)}$and the lower alpha bound $\alpha^{(-)}$ equation (214) is applied twice.

$$\beta_k = 0.5\left(1 + erf\left(\frac{\alpha_k^{(+)} - t_{RUL,k}}{\sigma_{RUL,k}\sqrt{2}}\right)\right) - 0.5\left(1 + erf\left(\frac{\alpha_k^{(-)} - t_{RUL,k}}{\sigma_{RUL,k}\sqrt{2}}\right)\right) \tag{215}$$

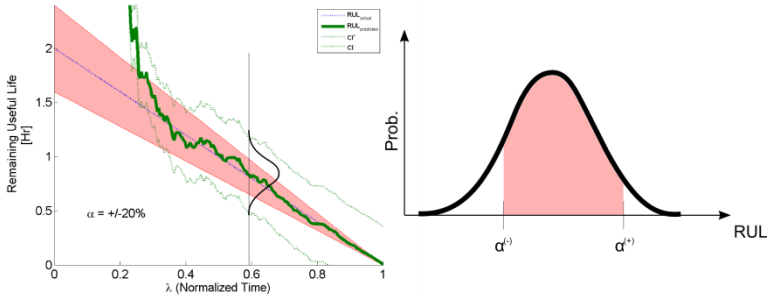


Figure 83: Illustration of the Beta calculation for a Gaussian distribution at a prediction 60% of the way to failure

The optimal score for a beta metric is unity at each prediction point. The minimum value for the beta metric is zero at each prediction point. For a time series of k predictions the beta metric is a vector with k entries. For the simulated data set the beta metric is shown in Figure 84.

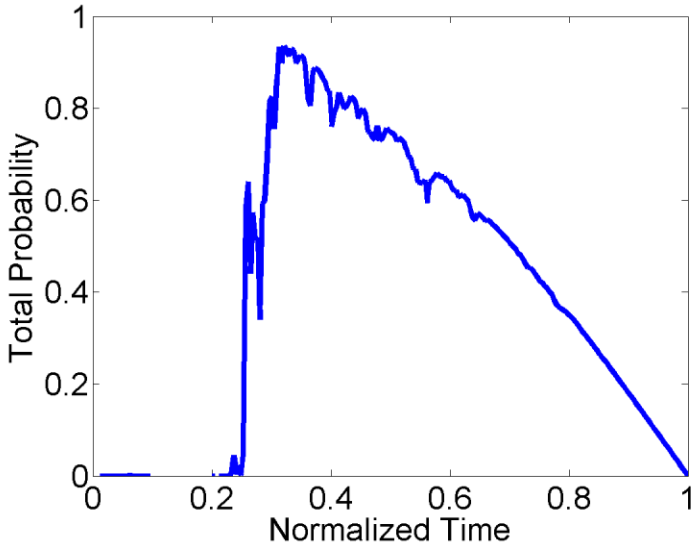


Figure 84: Beta metric for the simulated data set. Note the decrease in the metric score near the actual time of failure

6.3.2.2 Generic β metric

When the distribution describing uncertainty in the RUL prediction is a generic distribution as is the case in the particle filter a probability mass function is used to describe the distribution. Integrating a probability density function is straightforward using a simple integration technique for a discretely defined function. If the probability mass function is defined with a fine discretization of points, the rectangular rule for integration can be implemented.

$$\beta_k = \int_{-\alpha}^{+\alpha} PDF(x)dx \approx \sum_{i_\alpha=2}^{N_\alpha} \left(x_{i_\alpha} - x_{i_\alpha-1}\right) \frac{PMF\left(x_{i_\alpha}\right) - PMF\left(x_{i_\alpha-1}\right)}{2} \tag{216}$$

Where $i_\alpha = 1, \ldots, N_\alpha$ is the set of probability mass points inside the alpha bounds and PMF is the value of the probability mass function at each discrete point

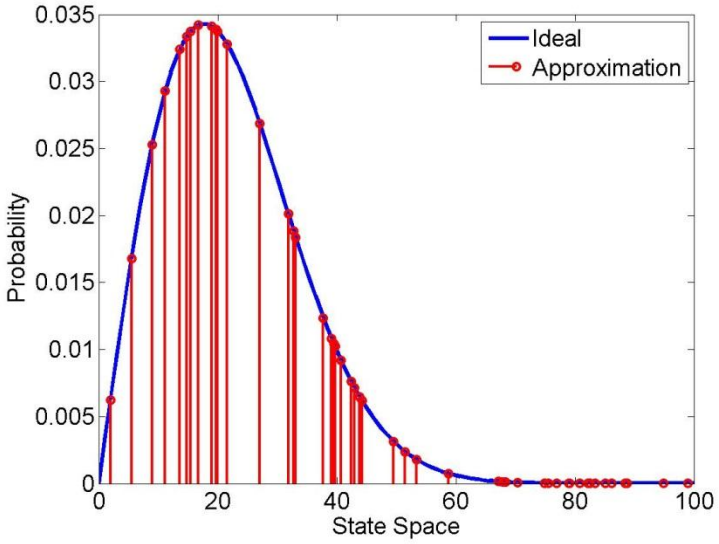


Figure 85: Numeric integration could be used to evaluate the area between two alpha bounds defined by the probability mass function denoted as the approximation to the blue ideal curve. This situation is common in a particle filter implementation base PHM algorithm.

6.3.3 Relative Accuracy

Relative accuracy is defined as

$$RA_k = 1 - \left| \frac{RUL_{actual} - RUL_{predicted}}{RUL_{actual}} \right| \tag{217}$$

A perfect relative accuracy score is one, and the worst score is zero. In the case of a prediction that severely deviates from the actual RUL the metric can be negative. If a negative value for the metric is calculated the metric is reported as zero. For a time series of k predictions the RA metric is a vector with k entries. The relative accuracy score for the simulated data set is shown in

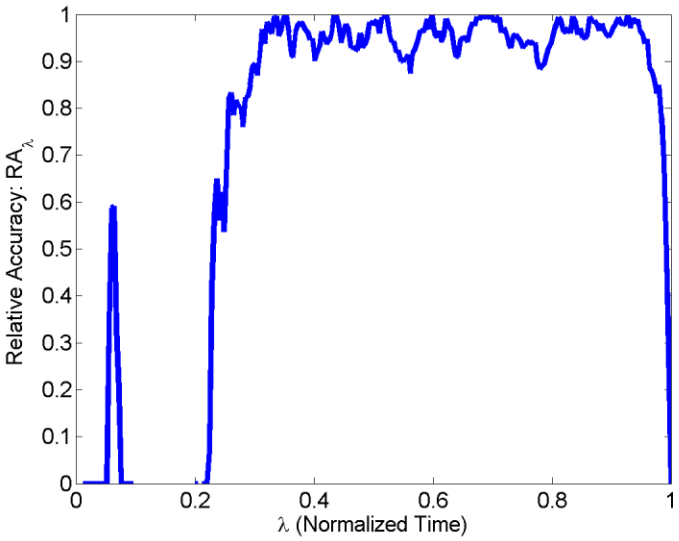


Figure 86: Relative accuracy for the simulated PHM predictions. Note the sharp drop off in relative accuracy towards the end of life.

6.3.4 Cost Function

To facilitate comparisons of different data sets/PHM algorithm results using a single number a cost metric is created. The cost metric is arbitrary and user specific. In this paper we took an equal weighting of accuracy and precision to determine the performance of the algorithm.

$$J = \frac{1}{N}\sum_{k}\left(C_\beta * \beta_k + C_{RA} * RA_k\right) \tag{218}$$

Where, β and RA are vectors containing a value between zero and one for each prediction time step and N is the total number of predictions. C_β and C_{RA} are weights that sum to a value of unity. In this case $C_\beta = C_{RA} = 0.5$. The resulting performance function for the hindcaste was 0.443.

6.4 Conclusion

In this section the off line performance metrics that will be used to quantify performance of presented prognostic algorithms was defined and their use was demonstrated on a synthetic data set. A remaining useful life plot succinctly summarizes the performance of algorithms to be quantified as a function of time. The alpha lambda metric was helps quantify the precision of RUL predictions. The beta metric quantifies the accuracy of RUL predictions. Relative accuracy is the preferred method for summarizing the error between actual and predicted RUL. To assign a single number to the performance of an algorithm, a cost function was defined as a weighted average of the beta metric and relative accuracy metric. When used together, the PHM metrics provide a method for quantitative comparison between competing algorithm implementations.

7 Prognostics with Resistance Spectroscopy for Grid Array Packaging

Leading indicators of failure have been developed based on high-frequency characteristics, and system-transfer function derived from resistance spectroscopy measurements during shock loadings. The technique is intended for condition monitoring in high reliability applications where the knowledge of impending failure is critical and the risks in terms of loss-of-functionality are too high to bear. Previously, resistance spectroscopy measurements [Constable 1992] have been used during thermal cycling tests to monitor damage progression due to thermo-mechanical stresses. The development of resistance spectroscopy based damage pre-cursors for prognostication under shock and vibration is new. In this section, the high-frequency characteristics, and system transfer function based on resistance spectroscopy measurements have been correlated with the damage progression in electronics during shock. Packages being examined include ceramic area-array packages. Second level interconnect technologies examined include copper-reinforced solder column, SAC305 solder ball, and 90Pb10Sn high-lead solder ball. Assemblies have been subjected to 1500g, 0.5 ms pulse [JESD-B2111]. Continuity has been monitored in-situ during the shock test for identification of part-failure. Models for healthy and damaged packages have been developed based on package characteristics. Data presented shows that high-frequency characteristics and system-transfer characteristics based on resistance spectroscopy measurements can be used for condition-monitoring, damage initiation and progression in electronic systems. A positive prognostic distance has been demonstrated for each of the interconnect technologies tested.

7.1 Test Vehicle

Test board A, was manufactured in four different configurations. The boards have daisy-chained ceramic packages with 400 I/O each. Each test board has 8-packages on one side of the test board. All packages on a test-assembly have the same interconnect type. Interconnect technologies studied include, copper-reinforced solder column grid array (CCGA), Eutectic tin-lead solder (63Sn37Pb), high lead solder (Sn10Pb90) and SAC305 solder (Sn3Ag0.5Cu). Table 8 shows the package parameters for test board A. A representative board with each package numbered U1 – U8 is shown in Figure 87.

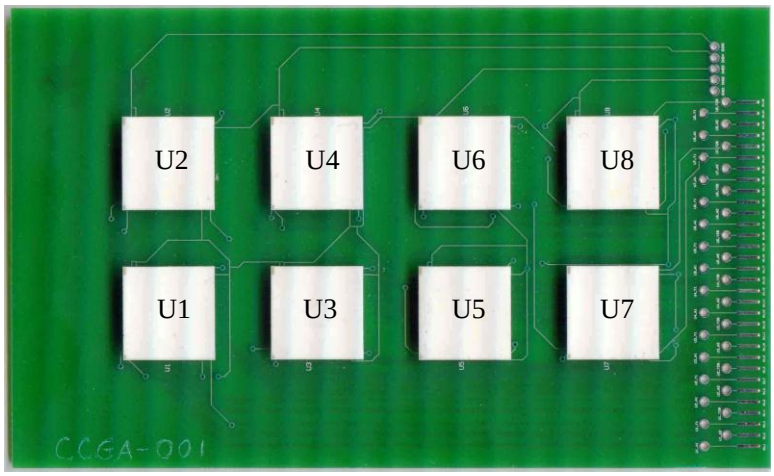


Figure 87: Test Vehicle

Table 8: Package Architectures for Test Board A

Parameter	CCGA Solder	CBGA 63Sn37Pb	CBGA Sn10Pb90	CBGA Sn3Ag0.5C
Length (mm)	21	21	21	21
Width (mm)	21	21	21	21
Thickness (mm)	2.4	2.4	2.4	2.4

I/O	400	400	400	400
Pitch (mm)	1	1	1	1
Ball Dia (mm)	0.6	0.6	0.6	0.6
Joint Height	2 mm	0.6 mm	0.6 mm	0.6 mm

7.2 Resistance Spectroscopy Approach

Leading indicators of damage have been developed based on resistance spectroscopy and impedance response phase shift of interconnects. The interconnect impedance response has been measured as a function of input signal frequency during shock and vibration. Interconnect resistance change can be represented by the equation,

$$R = \rho \frac{L}{A} \tag{219}$$

Where R is the resistance, ρ is the material property resistivity, and L and A are the length and cross sectional area respectively.

$$\frac{dR}{R} = \frac{d\rho}{\rho} + \frac{dL}{L} - \frac{dA}{A} \tag{220}$$

$$\frac{dR}{R} = \frac{d\rho}{\rho} + \varepsilon_a(1 + 2\nu) \tag{221}$$

The temperature coefficient of resistivity, is a function of temperature. Since the primary focus is on resistance spectroscopy measurements and their correlation with damage, the resistance change due to temperature has removed from the measurements. The temperature during damage initiation and progression has been held constant during experiments. The resistance change can thus be attributed to damage in the interconnects.

$$\frac{dR}{R} = \frac{dR_{temp}}{R} + \frac{dR_{damage}}{R} = \frac{d\rho}{\rho} + \varepsilon_a(1+2\nu) \tag{222}$$

$$\frac{dR_{damage}}{R} = \varepsilon_a(1+2\nu) \tag{223}$$

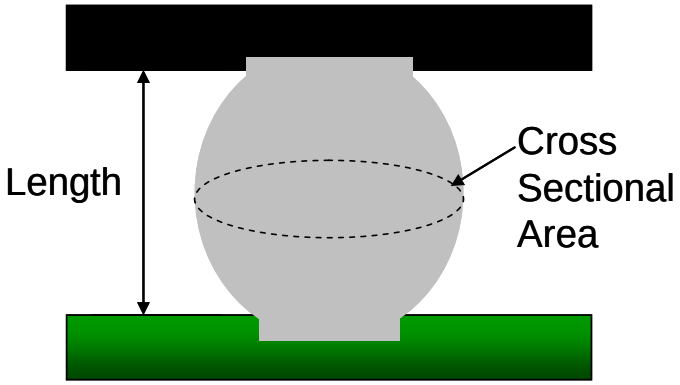


Figure 88: Schematic representation of solder ball denoting length and cross sectional area

Solder interconnects may be strained during shock and vibration. Homologous temperature, which expresses the temperature of a material as a fraction of its melting point temperature using the Kelvin scale, is in the neighborhood of 0.5 times its melting temperature at room temperature or normal equipment operating temperatures for several solder alloys. In electronics applications, where circuits typically operate over a –55°C-+125°C range, solder may be operating at 0.5-0.8 T_{melt}. Shock loading in electronic interconnects at equipment operating temperature will result in plastic deformation, residual strain and eventual failure of the interconnects. Resistance spectroscopy has been used to monitor the frequency dependent impedance response of the interconnects for monitoring damage initiation and progression.

7.3 The Effect of Shear Strain on Change in Resistance

It has been hypothesized that when dynamic loadings, such as drop and vibration, bend a printed circuit board the strains imparted on the solder joints are predominantly axial [Darveaux 2006, Chong 2006, Lall 2007b]. The modulus of rigidity of the package, shown in black in

Figure 89, is much higher than the equivalent modulus of rigidity of the layer of solder joints. Therefore when a curvature is applied to the printed circuit board, shown in green, the solder joints have the least rigidity and are stretched in the axial direction by the applied curvature of the PCB.

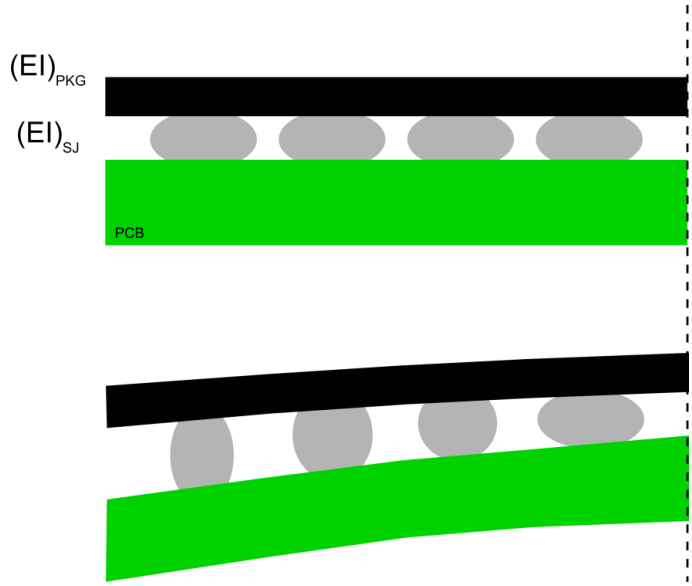


Figure 89: Hypothesized effect of PCB bending on solder joint interconnects

Resistance of a cylindrical conductor, as described in equation (219), is a function of the material characteristics and the geometry of the conductor. If it can be assumed that changes in the materials resistivity are negligible, then only changes in geometry affect the resistance of a conductor. To understand the effect of shear and axial strain on a solder joint with a more complicated geometry, a finite element simulation of a single solder joint was created. Separately axial strains and shear strains were applied to the joint and the change in resistance was calculated for the deformed solder joint geometry. The simulation used linear elastic material properties for the solder, which is known to not be realistic, therefore only a qualitative comparison between the axial strain case and the shear strain case will be examined.

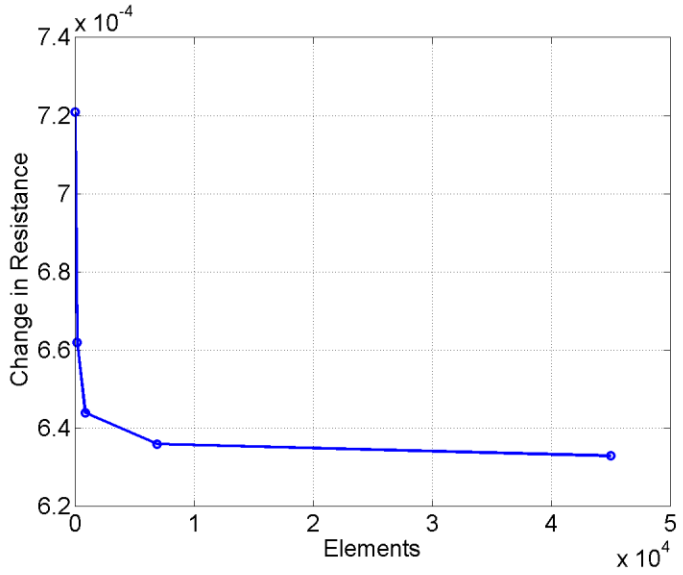


Figure 90: Mesh convergence study for the solder joint geometry under an axial loading

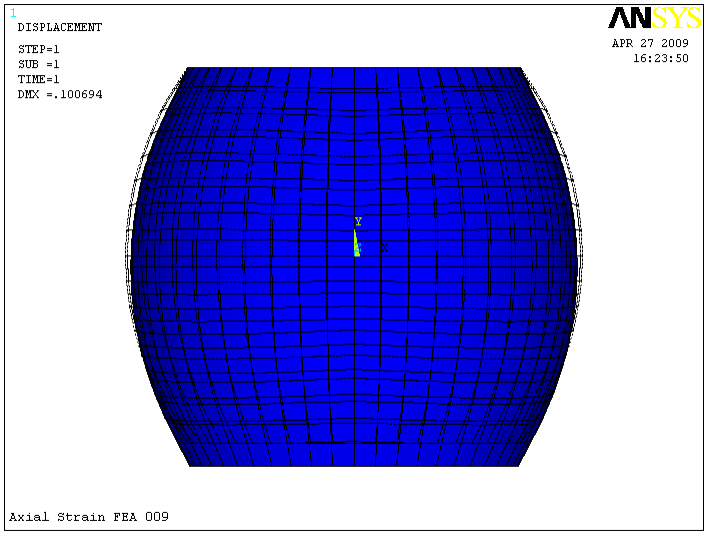


Figure 91: Axial strain case with the top of the solder joint constrained and the bottom of the solder joint is displaced downward. The grid represents the unreformed shape, and the solid represents the deformed shape

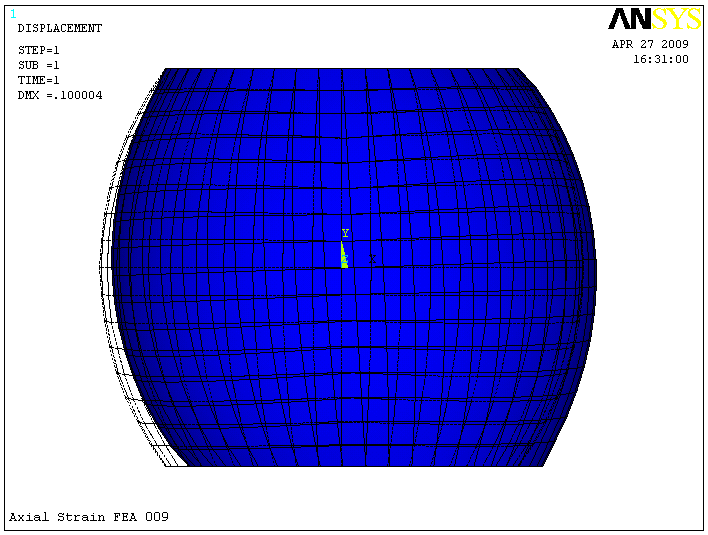


Figure 92: Shear strain case with the top of the solder joint constrained and the bottom of the solder joint is displaced to the right. The grid represents the unreformed shape, and the solid represents the deformed shape

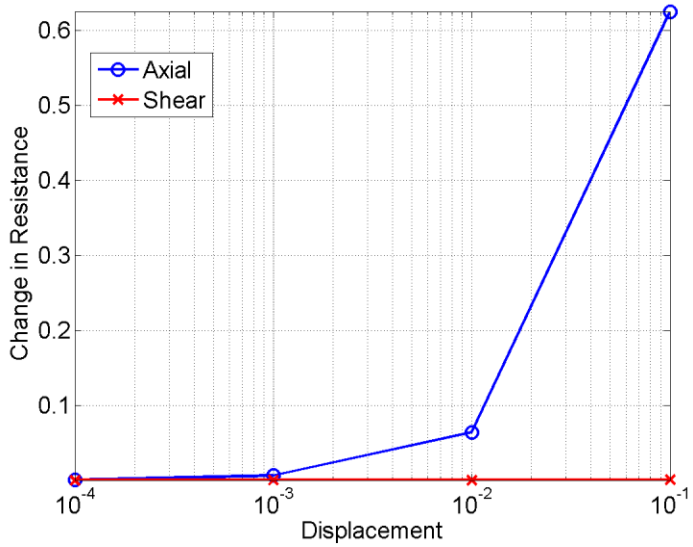


Figure 93: Comparison of trends in the calculated change of resistance for the two simulations

In Figure 90 the mesh convergence study for the solder joint model is shown. Figure 91 and Figure 92 shows the deformed states of the solder joint for the axial and shear stain simulations respectively. Figure 93 highlights that the contribution to change in resistance from a solder joint is predominantly from axial strains. In general the change in resistance from shear

strain is many orders of magnitude less than the contribution from axial strain, but not exactly zero. Introductory mechanics texts [Hibbeler 2004] often describe pure shear strain as a process that causes no volume change in a material. The small but non zero change in resistance seen in the above simulation is believed to be the result of not applying a pure shear boundary condition.

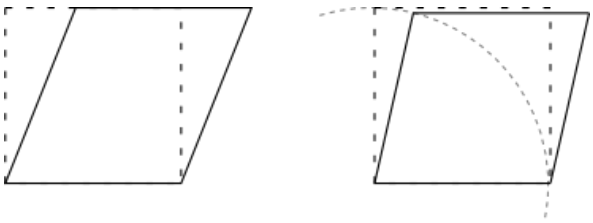


Figure 94: The simulated state of shear, shown on the left, is not a pure state of shear and adds small amounts of axial strain to the simulation. A pure state of shear creates no change in length of a strain element and is shown on the right. The pure state of shear is difficult to apply to a solder joint geometry.

7.4 Phase Sensitive Detection

Phase sensitive detection was described in detail in a previous section and has been used for measurement of the resistance change and phase change. . For an electronic package with an initial resistance of 2.5Ω, and a residual strain of 1000με, a change in resistance of 1.1mΩ could be expected. Resistance spectroscopy has been previously used for thermal fatigue damage of solder interconnects [Constable 1992, Constable 2001, Lizzul 1994, Prassana 1995]. The use of the technique for prognostication of shock and vibration induced damage is new. Even though commercial multimeters are available with resolution near 1mΩ, many factors such as contact resistance and metals relatively high thermal coefficient of resistance can skew a measurement in the 1mΩ range. Resistance spectroscopy is a technique that can be used to measure changes in resistance well below the 1mΩ level. An AC Wheatstone bridge has been used in conjunction with a lock in amplifier that measures the magnitude and phase shift of the signal. The work

presented in this paper utilizes the phase shift of the signal output from the AC bridge instead of the magnitude of the signal. The setup is shown in Figure 95.

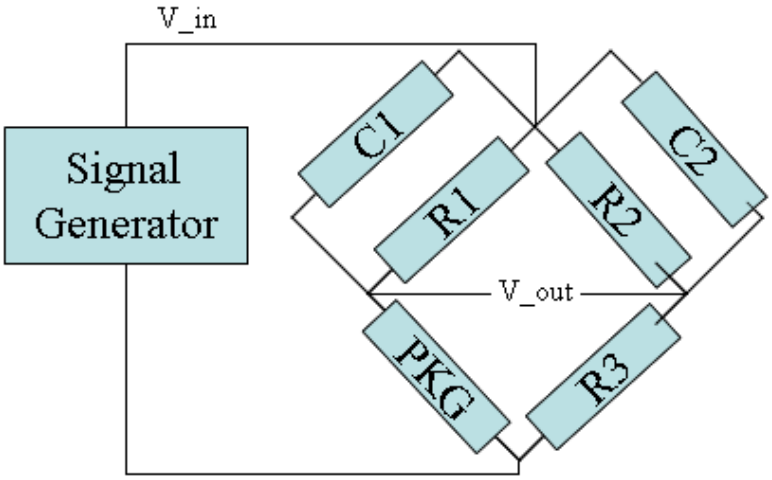


Figure 95: Wheatstone Bridge with capacitors C1, C2 and resistors R1, R2, R3, PKG

Two capacitors have been added to the bridge balancing resistors and the constant voltage source replaced with a signal generator. The daisy chained package has been denoted by the acronym, PKG in Figure 95. A change in package resistance will produce a change in the magnitude and phase shift of the output signal. The transfer function for the circuit can be represented by,

$$V_{out} = V_{in} \frac{Z_3 Z_1 - Z_4 Z_2}{(Z_2 + Z_3)(Z_1 + Z_4)} \tag{224}$$

The driving frequency of the input signal has been swept in the frequency range of 1 Hz to 1 MHz to measure damage initiation and progression due to shock and vibration. The bridge response has been measured for both magnitude and It is believed that data at the high end of the frequency range starts to be dominated by high frequency characteristics such as the skin effect and deviates from the theoretically predicted model.

7.5 Experimental Setup

The test assemblies were mounted face down on a Lansmont Model 23 drop tower and subject to a 0.5 millisecond, 1500G impact pulse in accordance with JEDEC standard JESD-B211. Continuity for damage detection was done during the drops. Electrical continuity was monitored at 10 million samples per second during the test. High speed digital video for use with digital image correlation software was also recorded during the drop test. The boards were subjected to resistance spectroscopy including both magnitude and phase-shift measurements between drops. Phase shift measurements were repeated between drops until all packages on the board failed. The packages show varying degrees of damage when electrical continuity failures occur. A metric has been created to quantify the damage sustained by the packages. Figure 97 shows how the number of open events, multiplied by the severity of the event to arrive at a point system quantifying the severity of the failure. Each package is interrogated by the resistance spectroscopy technique individually. Switching has been used to cycle through all the packages on the test board.

Figure 96: Lansmont Model 23 Shock Test System

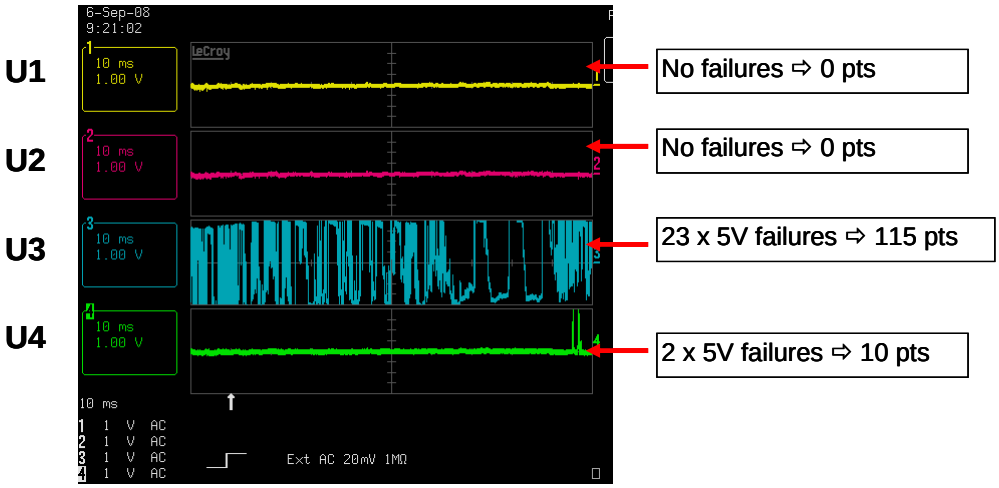


Figure 97: Failure Metric Calculation from Electrical Continuity.

7.6 Training Signal and Identification of Onset of Failure

Each measurement data-set has been processed for measured magnitude of the output signal and the phases shift with respect to the input signal. The phase shift of the output signal from the Wheatstone bridge at has been measured at 19 different frequencies. At each frequency a sample size of 200 data points has been recorded. Figure 98 shows the bode plot of the healthy package.

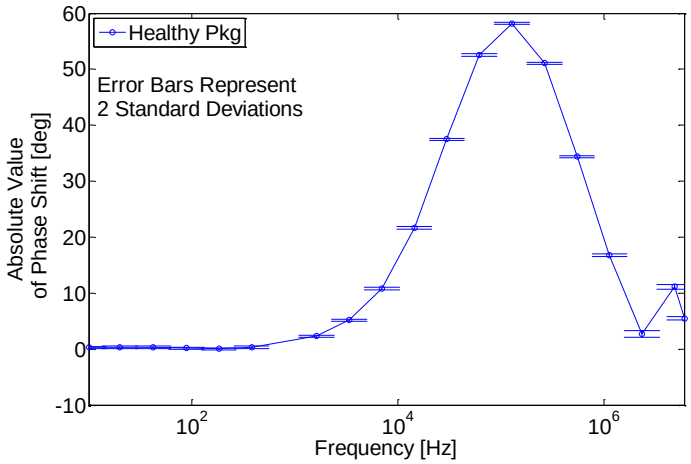


Figure 98: Phase shift of healthy package (Test Board-A, SAC305)

Bode plots have been constructed for each package on the test board after each drop and finite time increments during vibration. A training data-set has been constructed for healthy packages on the test assemblies. Variation in the transient data set for healthy assemblies and the deviation from the training data set in the damaged assemblies has been quantified using the t-statistic.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \tag{225}$$

$$DF = \frac{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^2}{\frac{\left(\frac{\sigma_1^2}{n_1^2}\right)^2}{n_1 - 1} + \frac{\left(\frac{\sigma_2^2}{n_2^2}\right)^2}{n_2 - 1}} \tag{226}$$

Where σ is the standard deviation, n is the number of samples in the data-set, $\bar{X}$ is the data-set mean, subscripts 1 and 2 represent the two data-sets, and DF is the number of degrees of freedom. Figure 99 shows the repeatability of the phase shift measurement for healthy packages at different locations on the same board.

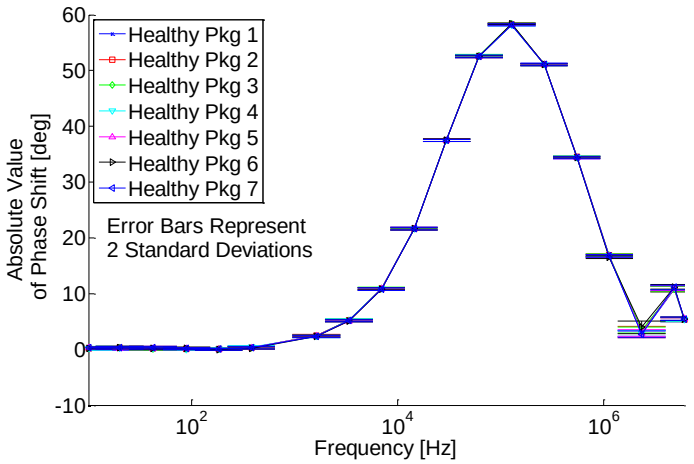


Figure 99: Repeatability of phase shift measurement on pristine healthy packages (Test Board-A, SAC alloy)

The data for the healthy pristine configuration of a package is used as a baseline for the confidence value calculation. The deviation in the magnitude and phase shift characteristic has been monitored from pristine state to a damaged state prior to failure, and eventual failure. Figure 100 shows the confidence value based on phase shift measurements of a package versus number of shock events in order to determine the statistical significance of the change in the phase shift of a package at a single frequency with respect to its original baseline healthy configuration. The confidence value degradation has been overlaid with the failure metric defined as the number of open events multiplied by the severity of the event to arrive at a point system quantifying the severity of the failure.

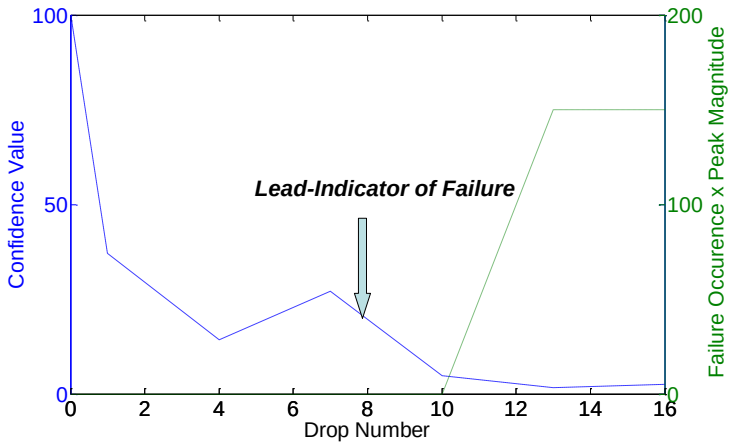


Figure 100: Confidence value as a lead indicator of failure during a drop test (Test Board-A, SAC alloy)

A higher failure metric value implies the number and severity of electrical opens is increasing. The failure metric indicates that the package failed electrical continuity in the neighborhood of the twelfth drop. Degradation in the confidence value precedes failure in the package interconnects. The correlation of the decrease in confidence value as a leading indicator of failure is not identical for all of the frequencies used in this experiment. In this plot only data measured at input signal frequency of 127kHz is displayed. In general it has been noticed that the

correlation between the degradation in confidence value and increase in resistance is better at higher input frequencies.

7.7 Prognostic Distance

A prognostic distance metric has been used for analyzing the damage proxy's ability for use as a leading indicator of failure. Statistical significance of phase shift with respect to the training signal has been used as one of the leading indicators of failure. Prognostic distance is defined as the lead-time between the indication of impending failure and electronic assembly failure. For shock-loads the prognostic distance has been measured in number of shock events. Impending failure is indicated by degradation in the confidence value that damage proxy in current assembly configuration is identical to that in the pristine undamaged assembly. Table 9, Table 10, Table 11 show the prognostic distances measured on various interconnect systems in a shock environment at input signal frequencies of 127 kHz, 2.33 MHz, 6MHz respectively. The prognostic distance is positive in all cases with no false positives measured in any configuration examined. The positive prognostic distance indicates that the damage proxy significantly precedes catastrophic failure of the package interconnects.

Table 9: Prognostic Distance (Test Board A, 127kHz)

Interconnect Type	Frequency	Threshold Confidence Value (%)	Number of Samples	Average Prognostic Distance (No. of Drops)	Std. Deviation (No. of Drops)	False Positives
SAC305	127kHz	90	3	6.33	3.06	0

63Sn37Pb	127kHz	90	4	24.75	8.62	0
90Pb10Sn	127kHz	95	6	18.67	6.02	0
Cu-CCGA	127kHz	30	4	41.5	9.47	0

Table 10: Prognostic Distance (Test Board A, 2.33MHz)

Interconnect Type	Frequency	Threshold Confidence Value (%)	Number of Samples	Average Prognostic Distance (No of Drops)	Std. Deviation (No of Drops)	False Positives
SAC305	2.33MHz	50	3	8.33	5.03	0
63Sn37Pb	2.33MHz	70	4	23.75	13.43	0
90Pb10Sn	2.33MHz	85	6	18.17	10.65	0
Cu-CCGA	2.33MHz	30	4	22.75	10.9	0

Table 11: Prognostic Distance (Test Board A, 6MHz)

Interconnect Type	Frequency	Threshold Confidence Value (%)	Number of Samples	Average Prognostic Distance (No of Drops)	Std. Deviation (No of Drops)	False Positives
SAC305	6 MHz	35	3	9	4	0
63Sn37Pb	6 MHz	20	4	18.75	4.65	0

90Pb10Sn	6 MHz	25	6	13.17	9.85	0
Cu-CCGA	6 MHz	30	4	43.5	23.39	0

Figure 101 shows the confidence value degradation preceding failure. Zero drops on the x-axis in Figure 101 indicates the time when failure of the electronic assembly was detected. The significant degradation in statistical confidence values indicates that the impending failure can be detected before failure of the electronic assembly. Figure 102 shows the repeatability of the confidence value degradation trend for various interconnect types including the SAC305, 63Sn37Pb, 90Pb10Sn, and Cu-CCGA.

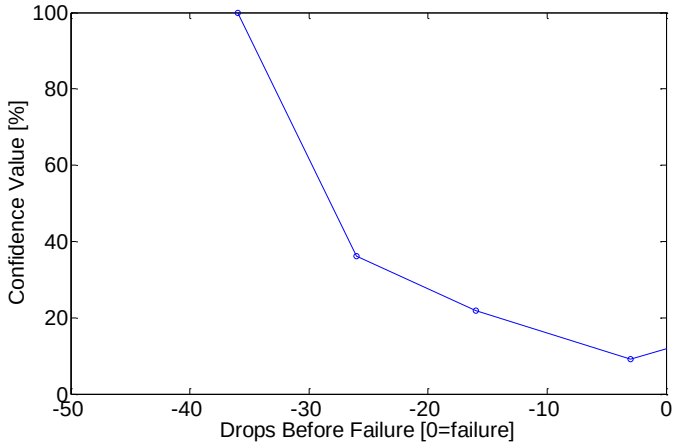


Figure 101: Degradation of confidence value during drop test (Test Board-A, CBGA Package U2, f = 6MHz)

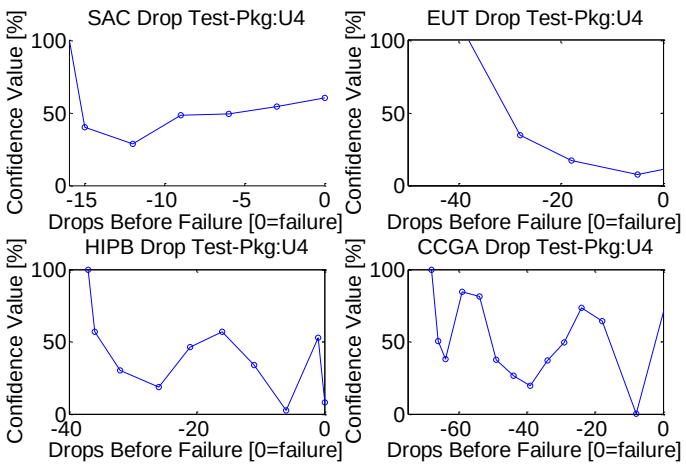


Figure 102: Degradation of confidence value during drop test (Test Board-A, f = 6MHz)

In each case, the degradation in confidence value significantly precedes failure exhibiting the ability of the damage proxy to serve as the leading indicator of failure. The difference in the degree of monotonic degradation is expected because of differences in the failure manifestation in the different interconnect systems. Figure 103 and Figure 104 show the consistency and repeatability of confidence value degradation trends for various packages of same interconnect type. The graphs show good repeatability.

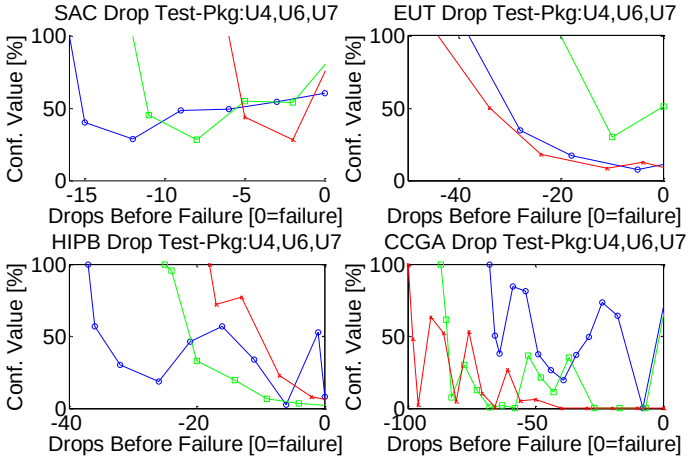


Figure 103: Degradation of confidence value during drop test of packages U4,U6 and U7 on for all interconnects. U4 is shown with blue circles, U5 is shown with green squares and U6 is shown with red crosses (Test Board-A, 127 kHz)

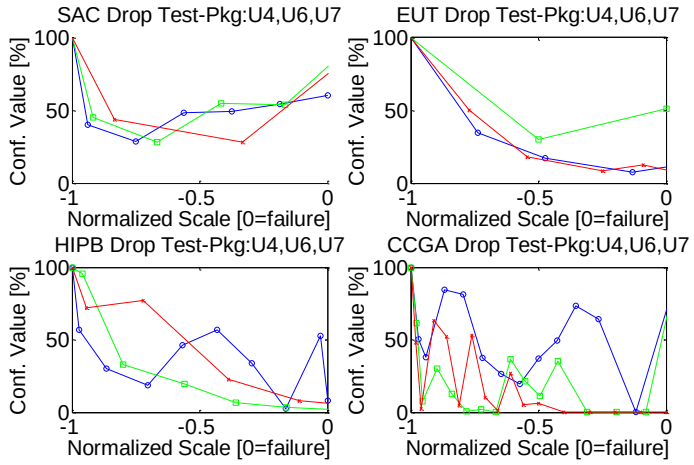


Figure 104: Degradation of confidence value during drop test of packages U4,U6 and U7 on a normalized scale for all interconnects. U4 is shown with blue circles, U5 is shown with green squares and U6 is shown with red crosses (Test Board-A, 6MHz)

A threshold value of 20-percent has been used to signify impending failure and calculation of prognostic distance. Figure 105 shows the confidence value threshold for SAC305 interconnects. The aggressive threshold value has been chosen to achieve statistical significance of difference to minimize occurrence of false positives. Figure 106 shows the distribution of prognostic distance for various input signal frequencies for SAC305 interconnects. Figure 106 illustrates the tradeoff between the prognostication distance and likelihood of a false positive. The prognostic distance is the highest for the 2.33 MHz input frequency, however the likelihood of false positives is also the largest. The prognostic distance is the smallest for the 6MHz input signal, however the propensity of false positives is also the smallest. For high reliability electronic applications false positives are completely unacceptable and therefore the system would be designed to avoid that type of error.

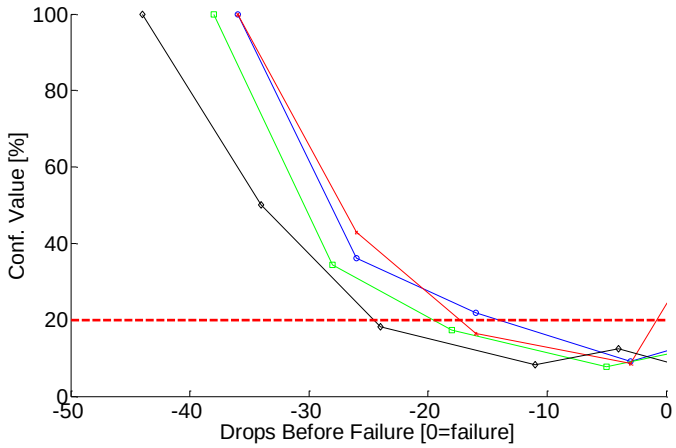


Figure 105: Method for determining prognostic distance using a threshold value shown in red. Each trace is an individual package. (Test Board-A, SAC305 alloy)

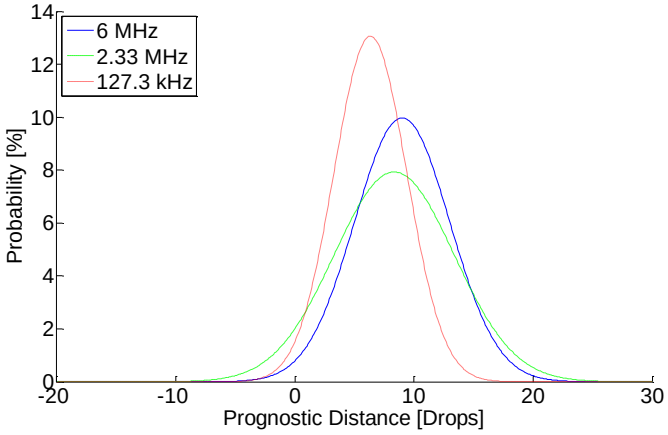


Figure 106: PDF of prognostic distance at varying frequencies. (Test Board-A, SAC alloy)

7.8 Failure Modes

Package failure modes have been verified with x-sectioning of the samples after failure. Several failure modes in the second interconnects have been observed. These include, trace cracking, solder interconnect failure in bulk or at the interfaces, and pad cratering. Figure 107 to Figure 110 show representative failures for all of the test board A interconnect types.

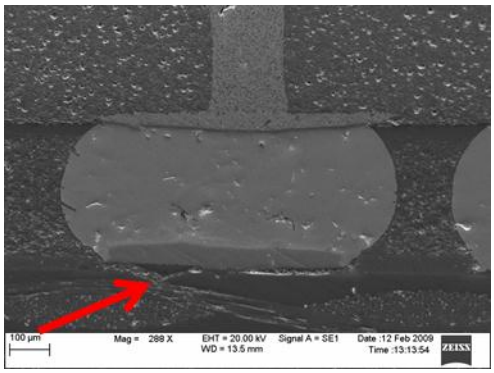


Figure 107: Damaged SAC interconnect (Test Board A)

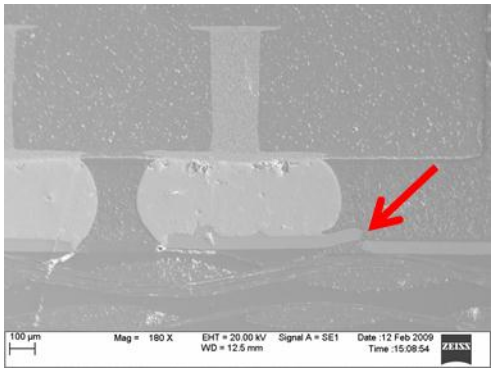


Figure 108: Broken trace on EUT interconnect (Test Board A)

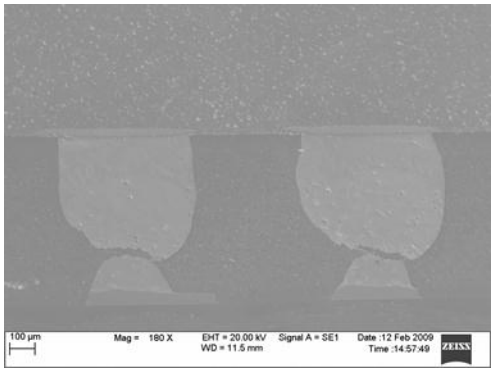


Figure 109: Cracks through HIPB interconnect (Test Board A)

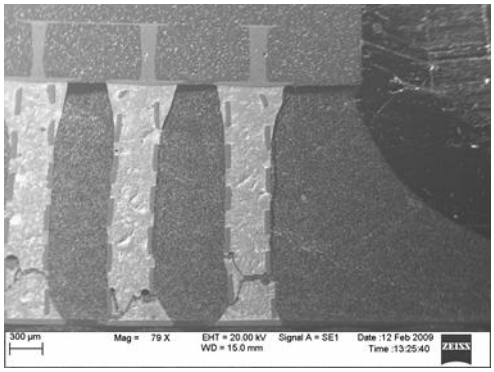


Figure 110: Failed CCGA interconnect (Test Board A)

7.9 Conclusion

A new technique has been developed for prognostication of second-level interconnect failures in electronic assemblies. The technique is based on damage pre-cursors derived from phase shift of the measured frequency response of the electronic assemblies using resistance spectroscopy measurements during shock loadings. Package architectures studied include ceramic area-array packages with multiple different second level interconnects including 63Sn37Pb, SAC305, copper reinforced solder columns, and 90Pb10Sn solder interconnects. Resistance spectroscopy based damage precursors have been measured during 1500g, 0.5ms shock and step-stress random vibration profiles. Data on eventual failure of the electronic assemblies has been gathered with high speed data-acquisition systems. Measurements of prognostic distance exhibit that the presented techniques can be used for early detection of impending failure in electronic assemblies. Prognostic distances have been quantified for the various interconnect structures examined and have been shown to be positive. The trade-off between prognostic distance and the propensity of false positives has also been examined and shown to have an inverse relationship. The failure modes have been studied in the packages after failure.

8 Recursive Least Squares (RLS) Filtering Based Prognostic Algorithm Implemented for Electrical Connectors

In this section a simple data driven prognostics algorithm is presented to predict failure in an electrical connector during an accelerated connector fretting test. While this work was originally intended for fielded electrical connectors, there has been interest in using the described techniques to shorten the duration of long running qualification tests. Specifically if a product is not going to pass qualification, advanced warning would allow the test to be canceled and save a considerable amount of time and money. Mechanical fretting increases the contact resistance in a connector and can ultimately lead to a failure in the electrical function of the system even if the mechanical connection is structurally sound.

8.1 Introduction

Mechanical fretting is a type of corrosion caused by the relative motion between two materials with a common interface. Fretting of electrical connectors has been studied in various forms, but the application of prognostics to predict a failure by mechanical fretting in electrical connectors is new. The mechanisms of fretting in electrical connectors similar to those used in this work have been studied [Flowers 2004, 2005, 2006, Xie 2009] in unpowered tests. Studies of connector fretting for powered tests have also been carried out [Angadi 2008, Polchow 2010, Fu 2010]. All tests for this work were performed at room temperature, but other authors have investigated the effects of connector fretting under a variety of different loading boundary conditions, and ambient temperatures/pressures [Daniel 2004, Lam 2006, Jedrzejczyk 2009, Swingler 2009, 2010]. Besides experimental investigations, analytical approaches aimed at modeling and predicting fretting in connectors has been performed [Bryant 1994, Zhang 2011,

Cartwright 2011]. Lastly this work tested a common tin coated connector, but various surface finishes have been shown to reduce the significance of fretting in electrical connectors [Swingler 2009, Fouvry 2011, Noel 2011]

8.2 Test Setup

In this test a connector system commonly found in desktop computers was investigated under the influence of vibration. The connector was a pin/spring configuration with 12 pins spaced 0.100” apart. Connectors were purchased from the Molex-Waldom corporation (Manufacturer Part Numbers, Header:22-23-2121, Housing: 22-01-2127). Traditionally the header portion of the connector is soldered to a circuit board and wires attached to the housing is routed from a distant source. To accelerate fretting in the connector a vibration input was applied to the header and the wires connected to the housing were mechanically tied off (constrained). This configuration is known to mimic the actual forces applied to connectors in fielded applications. The setup is also known to cause relative motion between the pin and spring inside the connector and therefore result in fretting corrosion and increased contact resistances.

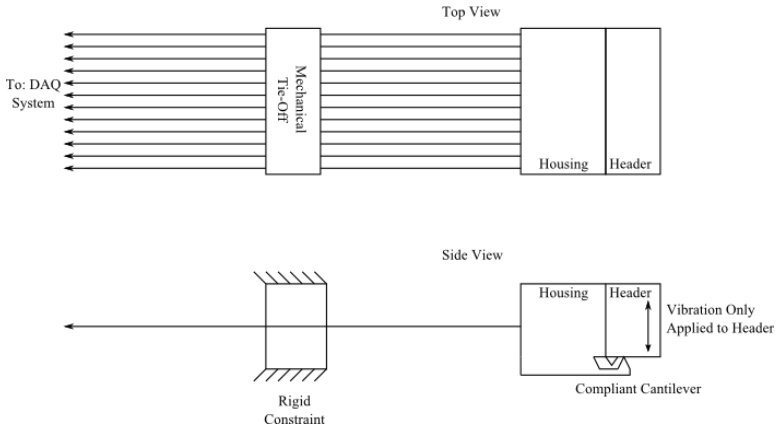


Figure 111: Diagram of Experimental Setup

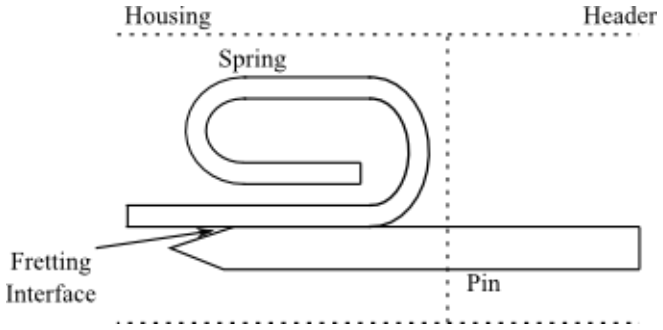


Figure 112: Diagram showing the internal configuration of the connector. Relative motion between the pin (rigidly attached to the header) and the spring (rigidly attached to the housing) results in fretting at the mating interface.

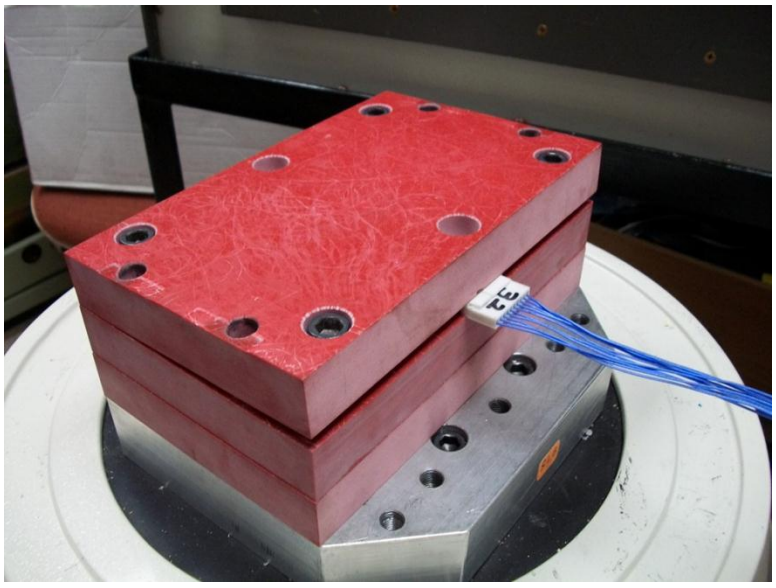


Figure 113: Fixture used to apply input to the connector header

Figure 114: Zoomed view of the fixture used to apply vibration inputs to the connector header

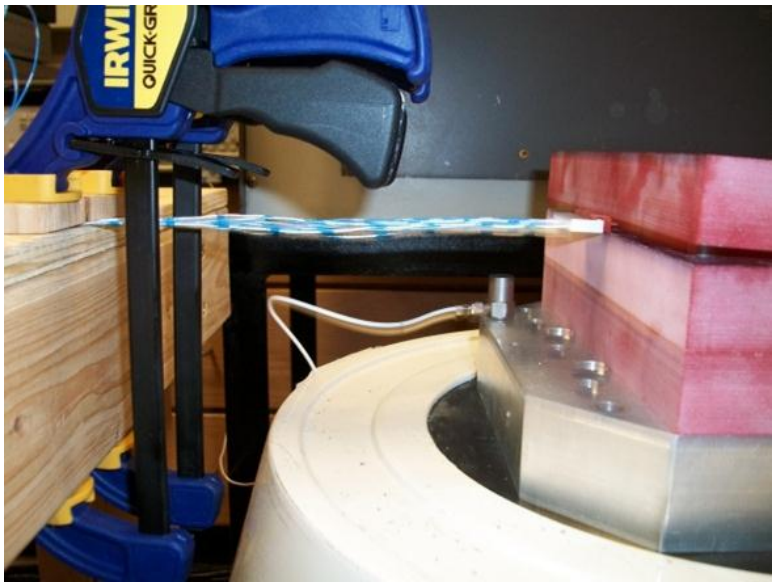


Figure 115: During testing the shaker table vibrates in the vertical direction (right side of picture). The motion of the shaker table has caused the image of the connector to blur. The rigid mechanical connection on the left shows no motion and did not blur in the photograph

8.3 Resistance Monitoring

Every second pin on the header was soldered together to allow sets of connections to be monitored using a two wire resistance measurement on an Agilent 34970A data logger. During vibration testing, resistance measurements were made on the first set of pins/springs every four seconds. The failure was defined as a rise in resistance of 0.3Ω. Since the data was very noisy a 1 minute moving average was applied to define the time of failure, which for the data shown in Figure 116 was at 5.2 hours.

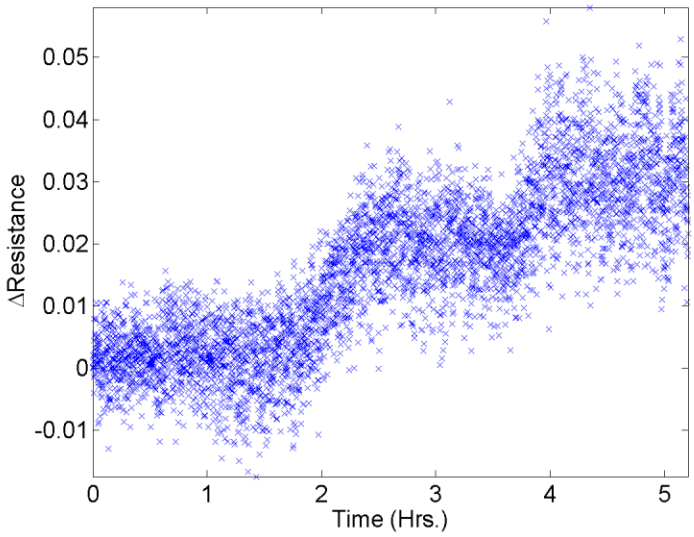


Figure 116: Time history of two wire resistance measurements and the increase in contact resistance caused by connector fretting

To ensure that the change in resistance was the result of fretting, and not an unrelated source the connector was disassembled and photographed with a light microscope.

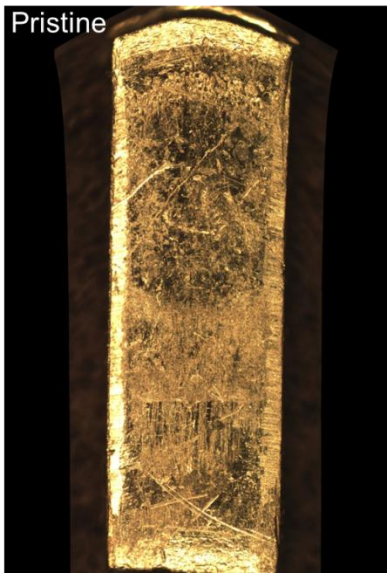


Figure 117: Contact surface of the spring before testing

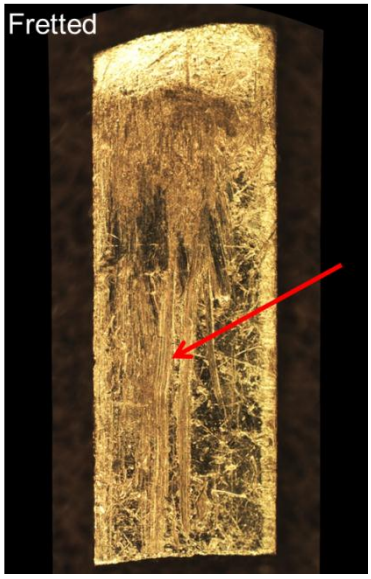


Figure 118: Deep grooves and asperities on a tested spring caused by relative motion between the spring and pin

The damage caused by testing is visible on the images of the spring part of the connector. The surfaces of the pin show similar damage as a result of vibration testing.

8.4 Prognostic Hindecaste

To simulate the ability to monitor and predict an increase in contact resistance that would cause the connector to fail the previously described data set was processed with a recursive least squares algorithm. Using the method of batch least squares a best fit line can be calculated. The R^2 value for the fit is 0.698 indicating a significant amount of un-modeled features in the data. Despite the less than desirable fit, it will be assumed that a linear model adequately describes the increase in contact resistance of the connector.

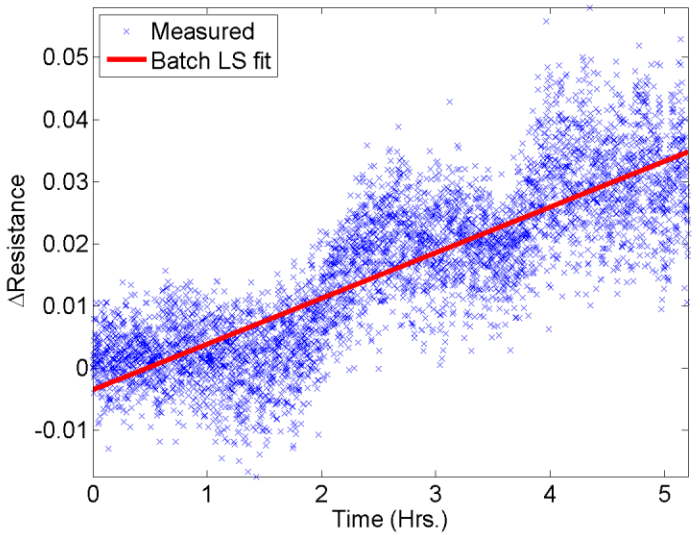


Figure 119: Batch least squares fit of a linear model to the experimentally observed data

The underlying principle of the prognostics approach is to recursively fit a line to the data available at the time of the prediction. By knowing the resistance at the prediction time, and an estimate of the change in the measurements with time, it is possible to prognosticate when the resistance of the connector will rise above the failure threshold of 0.3Ω. The two state recursive least squares equations described earlier are repeated for convenience.

$$\hat{\mathrm{x}}_{\mathrm{k}} = \mathrm{a}_0 + a_1 x \tag{227}$$

$$K_{1,k} = \frac{2(2k-1)}{k(k+1)} \tag{228}$$

$$\mathrm{K}_{2,k} = \frac{6}{k(k+1)T_s} \tag{229}$$

$$\bar{\mathrm{x}}_{\mathrm{k}} = \hat{x}_{k-1}^1 + \hat{x}_{k-1}^2 T_s \tag{230}$$

$$res_k = x_k^* - \bar{\mathrm{x}}_{\mathrm{k}} \tag{231}$$

$$\hat{x}_k^1 = \bar{\mathrm{x}}_{\mathrm{k}} + res_k K_{1,k} \tag{232}$$

$$\hat{x}_k^2 = \hat{x}_{k-1}^2 + res_k K_{2,k} \tag{233}$$

The RLS filter estimates the current change in resistance, and the slope of the resistance change measurement. At each time step that a measurement is taken, a prediction of remaining useful life will be made. RUL predictions at time step k are calculated using the state estimates from the RLS filter per equation (234).

$$RUL_k = \frac{x_f - \hat{x}_k^1}{\hat{x}_k^2} \tag{234}$$

Where x_f represents the failure threshold of 0.3Ω and $\hat{x}_k^1$ is the current estimate of the change in resistance in the connector at time k. The second state estimate, $\hat{x}_k^2$ is the first derivative, or slope of the best fit line.

The uncertainty surrounding the RUL prediction is based on a moving window of previous estimates of the slope of the recursively best fit line. The smallest and largest estimated slope for the last L measurements is propagated though equations (235) and (236) to bound the longest and shortest RUL that would be predicted at the current time based on the last L estimates.

$$RUL_{LOW} = \frac{x_f - \hat{x}_k^1}{min\left(\hat{x}_{(k-L):k}^2\right)} \tag{235}$$

$$RUL_{HIGH} = \frac{x_f - \hat{x}_k^1}{max\left(\hat{x}_{(k-L):k}^2\right)} \tag{236}$$

The uncertainty is reported as the difference between the high and low estimate for RUL

$$\sigma_{RUL} = RUL_{LOW} - RUL_{HIGH} \tag{237}$$

Along with the state estimates from the RLS filter, equations (234)-(237) are repeated for each new measurements. For this work the length of the window L was taken as 300 data points or 20 minutes. The result of RLS filtering is shown in Figure 120: Recursive least squares state

estimates. Initially the noisy data does not provide a significant trend to estimate meaningful state variables from. Due to the low estimated trend in the connector's resistance, RUL estimates greatly over predict the actual RUL.

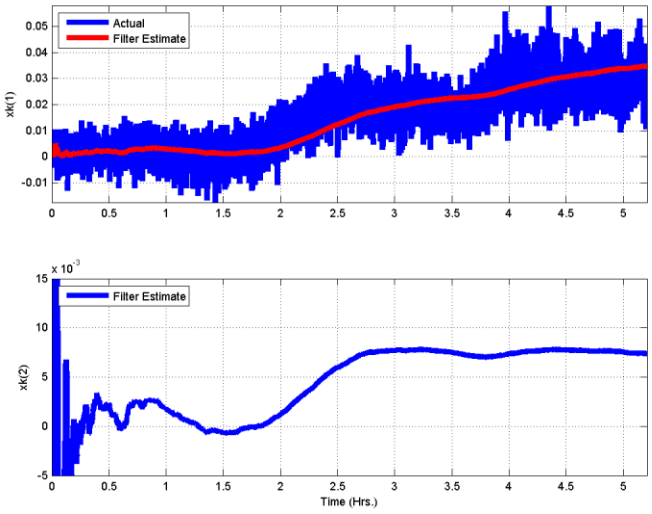


Figure 120: Recursive least squares state estimates

Approximately half way through the test the RLS filter converges on the true slope of the data set and RUL predictions converge to correct values. The accuracy of the PHM algorithm is summarized in Figure 121. At approximately the same time that the RLS filter converges on the correct slope for the best fit line, RUL predictions converge to values close to the actual RUL.

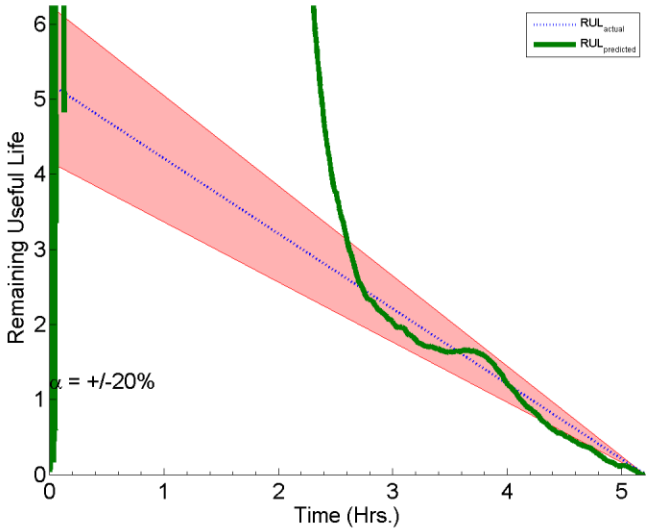


Figure 121: Remaining useful life predictions

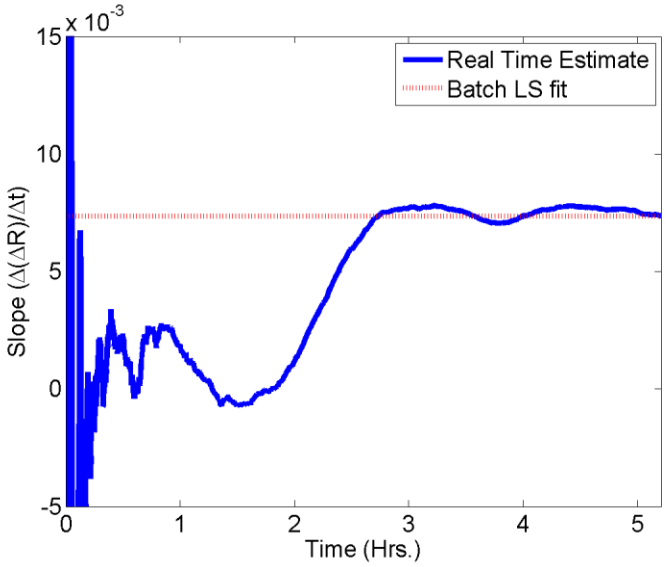


Figure 122: Comparison of batch least squares to recursive least squares estimates for the slope of the best fit line. Note that the batch and recursive estimate for the slope of the best fit line are identical for the full data set.

A comparison of the batch and least squares approaches are shown in Figure 122. Unfortunately the batch least squares method requires the entire data set for processing and is not a candidate for prognosticating failure before the connector has failed. Snapshots of the prediction process are shown in Figure 123 through Figure 126.

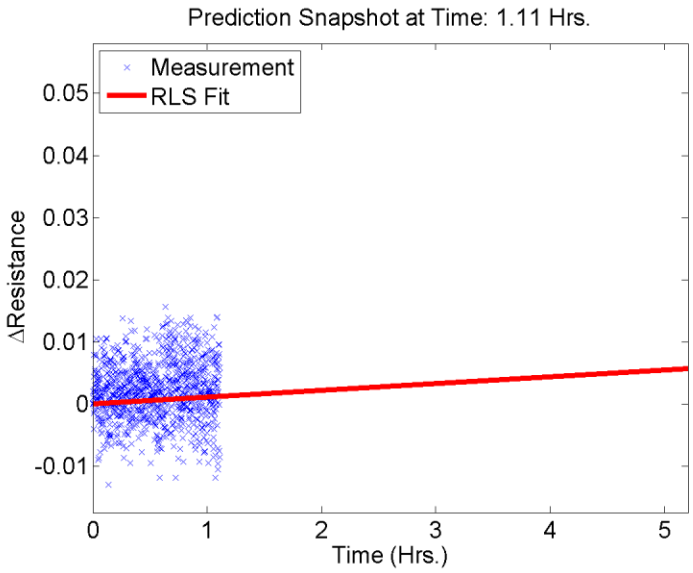


Figure 123: Prediction snapshot from early in the test. The lack of trend in the data results in an RUL prediction that is overly optimistic

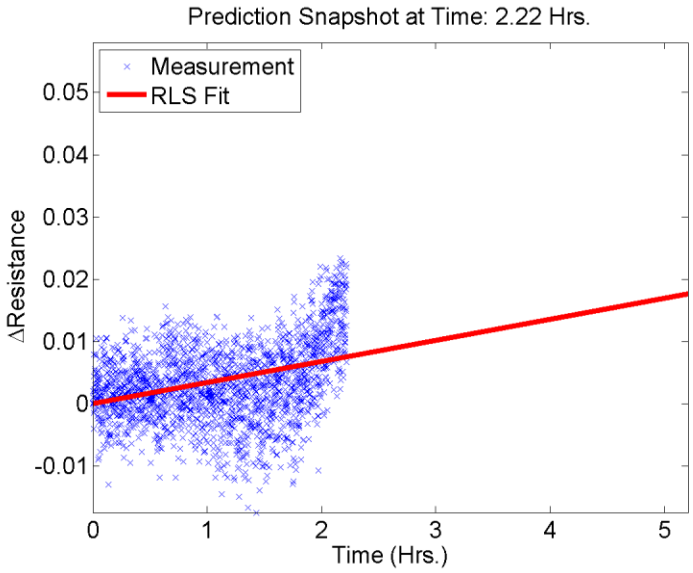


Figure 124: Further into the test RUL predictions are improving but still have not converged

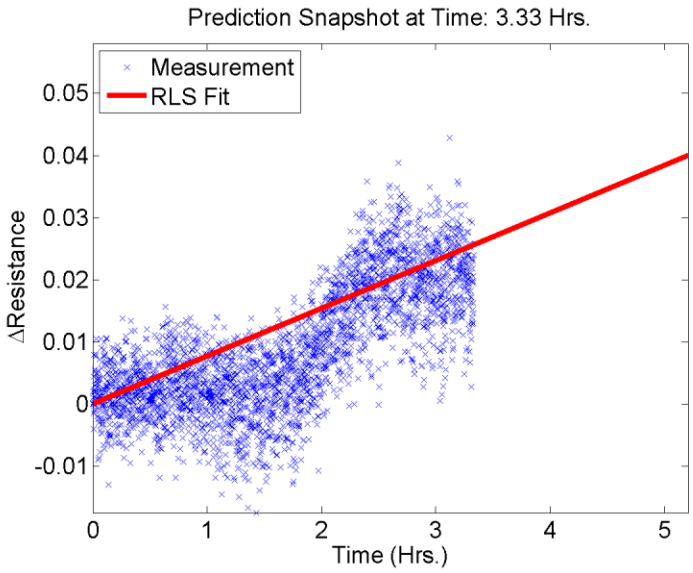


Figure 125: At this point in the test sufficient data is available to estimate the slope of a best fit line through the data

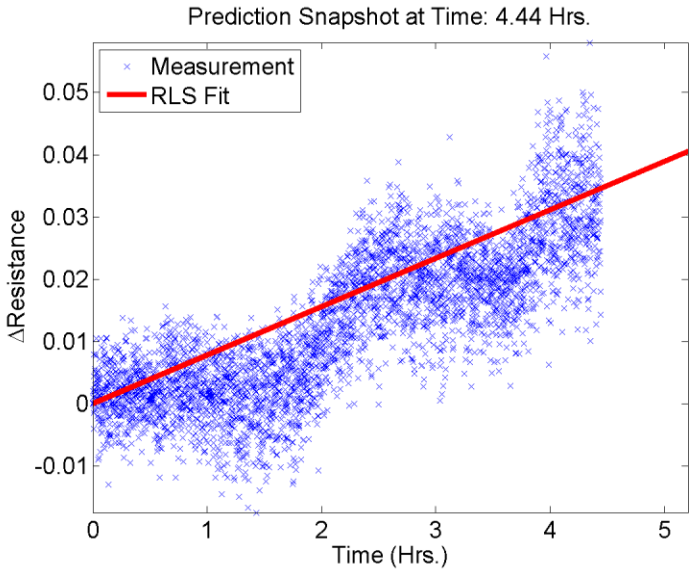


Figure 126: At the end of the test the estimate of the slope of the best fit line is nearly identical the RLS fit

Uncertainty surrounding the RUL predictions are summarized in Figure 127. The uncertainty is very large in the beginning of the test before the measured data has risen above the noise floor. As the algorithm converges, the uncertainty in the predictions quickly converges to a very small value. At the end of the test the uncertainty estimates are almost zero even though the

predictions are not absolutely correct. Figure 128 shows the time history of the predicted uncertainty without the RUL predictions. The uncertainty denoted on the y-axis varies from 10^5 hours (~11 years) to 10^-5 hours (<1 minute).

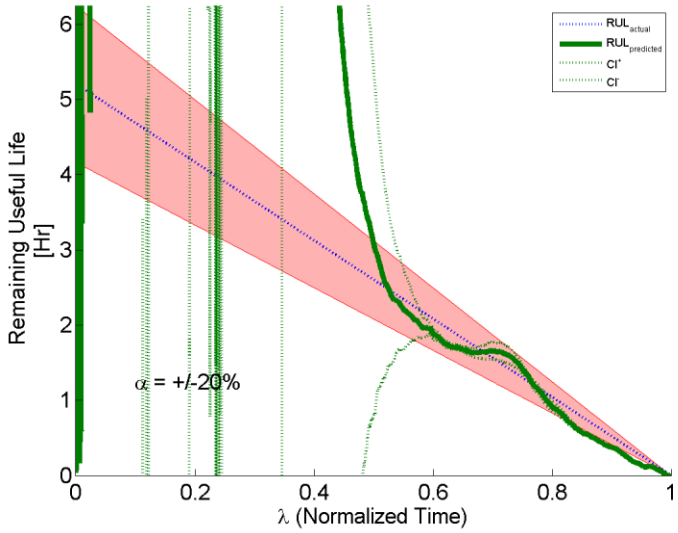


Figure 127: Alpha-lambda plot summarizing the uncertainty in the RUL predictions

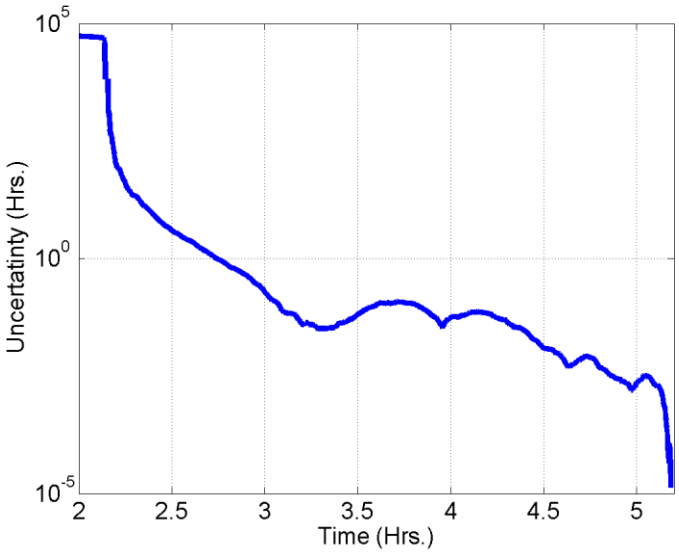


Figure 128: Time history of the uncertainty reported by the PHM algorithm. Note the log scale on the y-axis.

The beta and relative accuracy metrics are reported in Figure 129 and Figure 130 respectively. The binary performance of the beta metric as either very good or very bad is a

result of the very tight predictions of uncertainty. After converging about halfway through the test, the relative accuracy is good for this data set. The cost metric for this hindecaste was 0.622, where a perfect score would be zero.

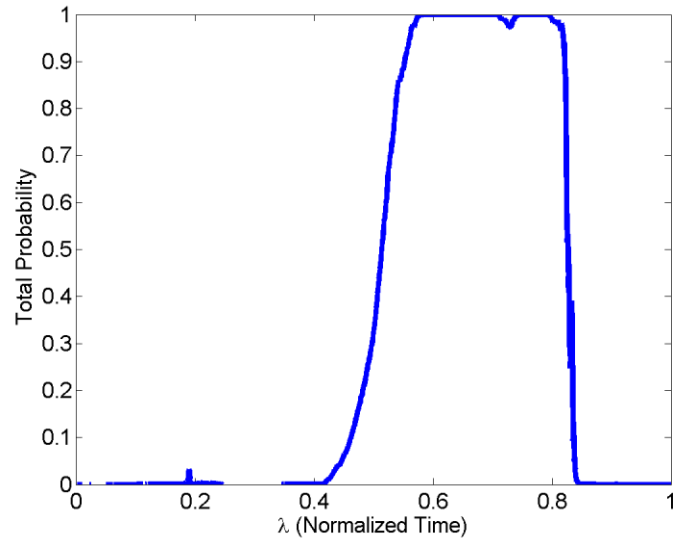


Figure 129: Beta metric calculation for the PHM hindecaste

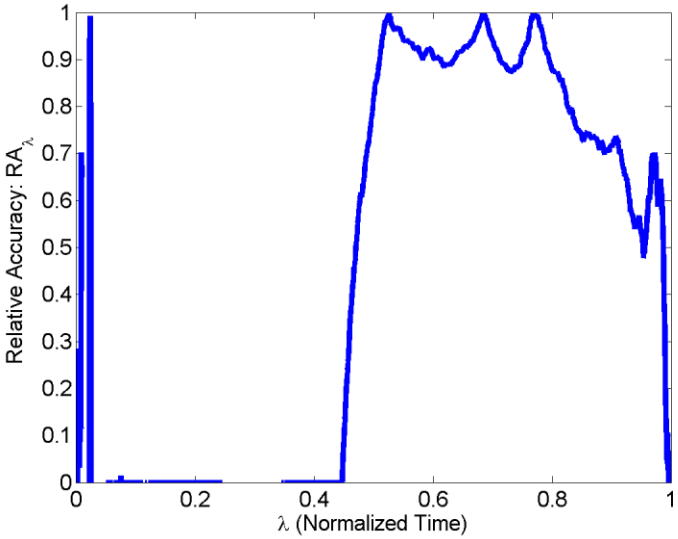


Figure 130: Relative accuracy metric for the PHM hindecaste

8.5 Conclusion

Admittedly the PHM algorithm implemented for this data set is very limited in its applicability. The model for failure degradation is completely data driven and implicitly assumes that usage conditions and inputs will not change. By its nature the RLS filter gives less and less emphasis to new measurements as the algorithm collects more data points and therefore would not tolerate a drastic increase or decrease in the rate of degradation. Regardless of its drawbacks this algorithm efficiently models the general procedure for PHM and how to apply and interpret the performance metrics.

9 Kalman Filter (KF) Based Prognostic Algorithm Implemented for BGA's

Structural damage to ball grid array (BGA) interconnects incurred during vibration testing has been monitored in the pre-failure space using resistance spectroscopy based state space vectors, rate of change of the state variable, and acceleration of the state variable. The technique is intended for condition monitoring in high reliability applications where the knowledge of impending failure is critical and the risks in terms of loss-of-functionality are too high to bear. Future state of the system has been estimated based on a second order Kalman Filter model and a Bayesian Framework. The measured state variable has been related to the underlying interconnect damage in the form of inelastic strain energy density. Performance of the prognostic health management (PHM) algorithm during the vibration test has been quantified using performance evaluation metrics. The methodology has been demonstrated on lead-free area-array electronic assemblies subjected to vibration. Model predictions have been correlated with experimental data. The presented approach is applicable to functional systems where corner interconnects in area-array packages may be often redundant. Prognostic metrics including alpha-lambda precision, β accuracy, and relative accuracy, have been used to assess the performance of the damage proxies. The presented approach enables the estimation of residual life based on level of risk averseness.

9.1 Introduction

Kalman filtering is a recursive algorithm that estimates the true state of a system based on noisy measurements [Zarchan 2000]. Use of Kalman Filtering for prognostication of electronic reliability based on the underlying damage mechanics is new. The Kalman filter has been utilized for this work since it is a robust tool for real time tracking of noisy signals, with a long

history of successful implementation. In this paper, a prognostic and health monitoring capability for electrical components based on changes in resistance has been presented. The presented PHM framework enables the estimation of remaining useful life in deployed electronics by interrogation of the system state and evolution of the state vector. The methodology has been demonstrated on area-array package board assemblies subjected to mechanical shock and vibration. Failure modeling of BGA interconnects is combined with Kalman filtering for plastic strain state estimation and a Bayesian framework for PHM. Prognostics metrics have been used to quantify the degree of uncertainty in the estimated remaining useful life.

9.2 Test Vehicle

A set of test boards with multiple package architectures were used for experimental measurements. The test board includes package architectures such as plastic ball-grid arrays, chip-array ball-grid arrays, tape-array ball-grid arrays, and flex-substrate ball-grid arrays . The experimental matrix has ball counts in the range of 64 to 676 I/O, pitch sizes are in the range of 0.5mm to 1mm, and package sizes are in the range of 6mm to 27mm. The package parameters of this board are shown in Table 12. Representative sample of the test board is shown in Figure 131.

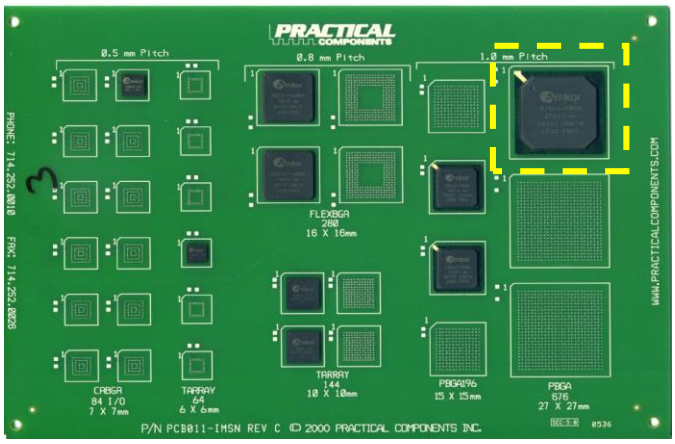


Figure 131: Test Board

Table 12: Package architectures on test board

	6 mm Tape Array	7 mm Chip Array	10 mm Tape array	15 mm PBGA	16 mm Flex BGA	27 mm PBGA
I/O	64	84	144	196	280	676
Pitch (mm)	0.5	0.5	0.8	1	0.8	1
Die Size (mm)	4	5.4	7	6.35	10	6.35
Substrate Thick (mm)	0.36	0.36	0.36	0.36	0.36	0.36
Pad Dia. (mm)	0.28	0.28	0.30	0.38	0.30	0.38
Substrate Pad	NSMD	NSMD	NSMD	SMD	NSMD	SMD
Ball Dia. (mm)	0.32	0.48	0.48	0.5	0.48	0.63

The test assemblies were subjected to vibration testing on a LDS Model V722 vibration table. A step stress profile was used to gradually ramp up the stress level to induce damage (Figure 132). The individual random stress profiles used in the step stress are shown in Figure

133. The next section will discuss how the transient response of a package during random vibration testing was monitored for a leading indicators of failure.

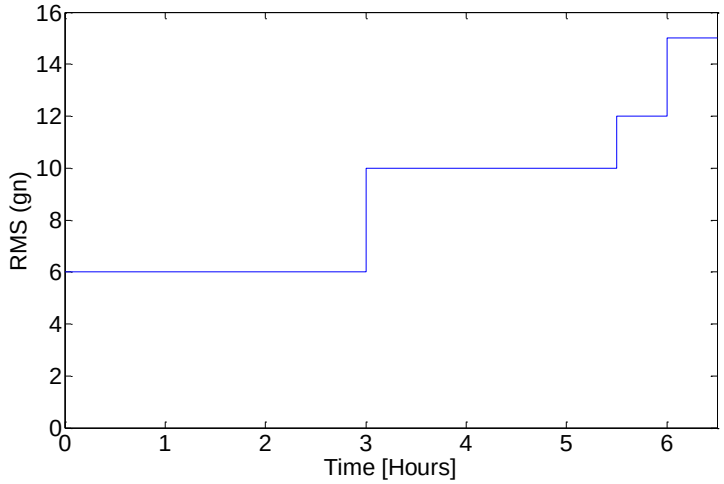


Figure 132: Step stress profile for vibration testing that fatigues interconnects to failure.

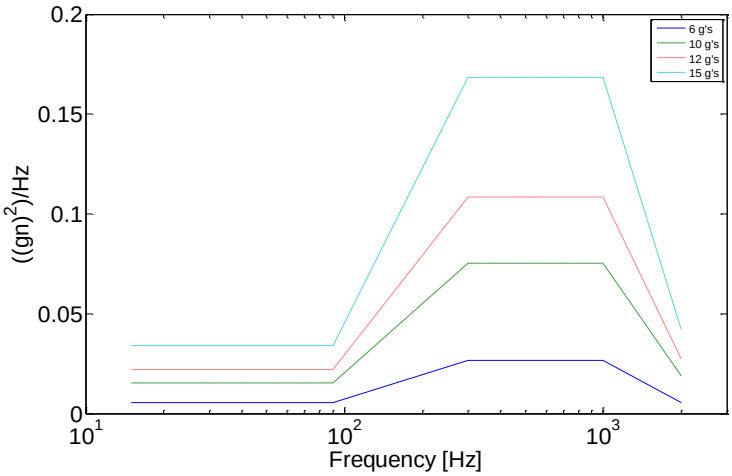


Figure 133: Random vibration profile at varying g levels corresponding to the step stress profile outlined in Figure 132

9.3 Transfer Function for Interconnect Strain to Resistance

The daisy chained resistance of a package was used as a leading indicator of failure in this paper. The observed history of the resistance of the package during vibration testing is shown in Figure 134.

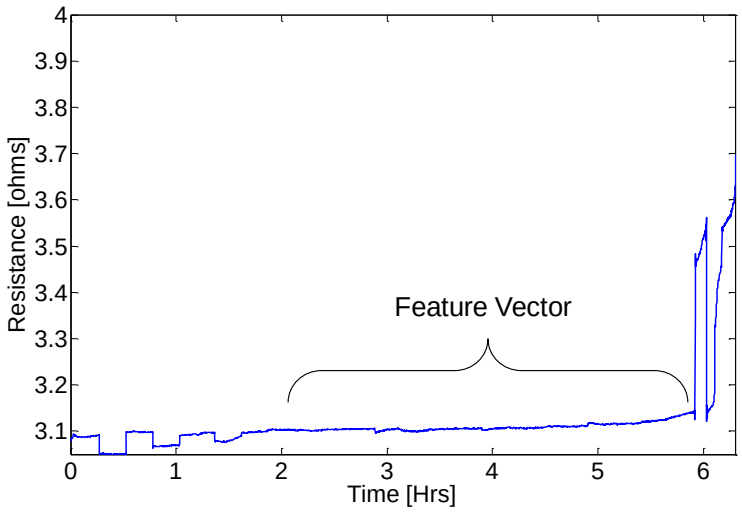


Figure 134: Raw resistance data. The data used as an input data vector is shown in the brackets

At approximately 5.8 hours the package experiences its first intermittent open event. In the following plots large resistance values have been truncated for clarity. The resistance of an open event of 300Ω or more makes it difficult to discuss mili-ohm changes on a plot. The resistance of the daisy chained package was recorded using an Agilent 34970A data acquisition unit with a two wire resistance measurement setup. Measurements were taken at a frequency of 0.2 Hz. Since data measurements were recorded every few seconds, but the test lasted for approximately 6 hours, this was deemed to be an effectively high sampling frequency to capture trends in the leading indicator of failure. Additional details quantifying the applicability of the measurement system for capturing intermittent events in advance of the traditional definition of failure can be found in [Lall 2009a-d].

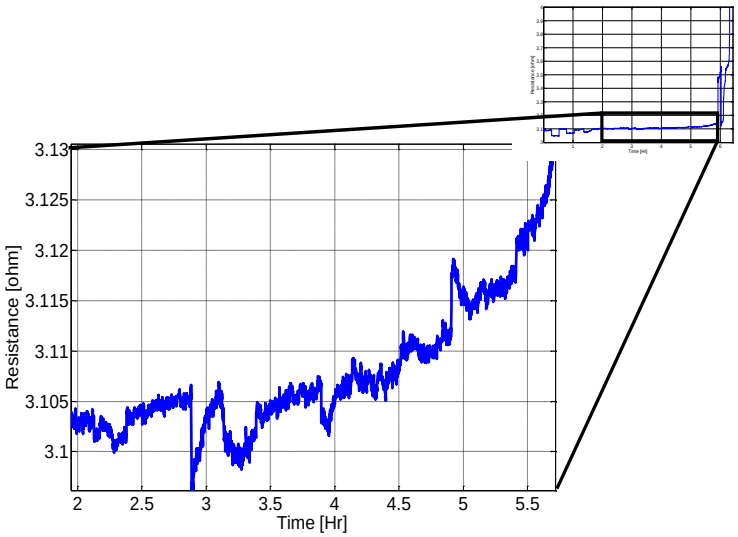


Figure 135: Zoomed view of resistance data between 2 hrs and failure

The failure criteria for resistance change outlined in industry standards JESD22-B103 [JEDEC 2006], and IPCSM785 [IPC 1992] for the number, duration, and severity of intermittent events is used as the definition of failure. It should be noted that the smaller step increases of 0.05 Ω during the first 90 minutes of the test are experimental noise which can be reproduced by motion of the system connections during shock and vibration. Resistance data two-hours after the initiation of the test till failure has been studied for the construction of a feature vector for identification of impending failure. A subset of the resistance data has been used since field data will often involve electronic assemblies with accrued damage and not involve pristine assemblies. Figure 135 shows a zoomed view of the input data highlighting the experimental noise between two hours and failure. The experimental noise is due in part to the challenges with overcoming the variance in contact resistance in the presence of transient dynamic motion in shock or steady-state vibration. Step changes in the resistance data can be seen at 2.8 and 4.9 hours respectively. However, the distinctive increase of about 25 mΩ during the vibration test is easily discernible even in the presence of experimental noise.

The change in resistance is attributed to change in geometry, since the resistivity of the solder interconnect is expected to stay constant. Change in trace geometry is the basis of operation for traditional strain gages and can be explained in a cylindrical conductor by $R = \rho L/A$, where R is the resistance of the conductor, ρ is the material property resistivity, L is length and A is the cross sectional area. By logarithmically differentiating both sides and assuming linear elastic properties, a relation between strain and resistance can be derived as $dR = R_0 \varepsilon_a (1+2\nu)$, where dR is the change in resistance, R_0 is the initial resistance, ε_a is the elastic axial strain and ν is the Poisson ratio. Since the material properties and geometry of a solder ball are non-linear, a finite element simulation (FEM) was used to map the change in resistance of an interconnect to the state of plastic strain that the interconnect was feeling. The simulation was implemented in ANSYS™ Version 12 using Anand's Viscoplasticity and VISCO107 elements. The Anand's constants used for the simulation are shown in Table 13

Table 13: Anand's Constants for SAC305

S_o	45.9 MPa
Q/K	7460 1/K
A	5.87e6 1/sec
ξ	2
M	0.0942
h_o	9350 MPa
n	0.015
a	1.5
s	58.3 MPa

Table 14: Undeformed geometry of solder ball

Parameter	Specification
Solder ball diameter (mm)	0.63
Solder ball land (mm, board and package)	0.45
Solder ball height after reflow (mm)	0.48

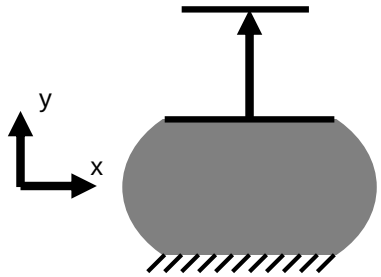


Figure 136: Constraints on solder ball for FEM simulation

Table 14 shows the dimensional parameters for the undeformed geometry of a typical solder ball based on the manufacture's data sheet. Previous studies have shown that tensile stress in the out-of-plane z-direction is the primary stress during the shock test in the solder interconnects [Darveaux 2006, Chong 2006, Lall 2009e]. The solder interconnect deformation during the shock test was simulated using non-linear finite elements by constraining the solder interconnect along the bottom of the joint and applying a displacement load on the top (Figure 136).

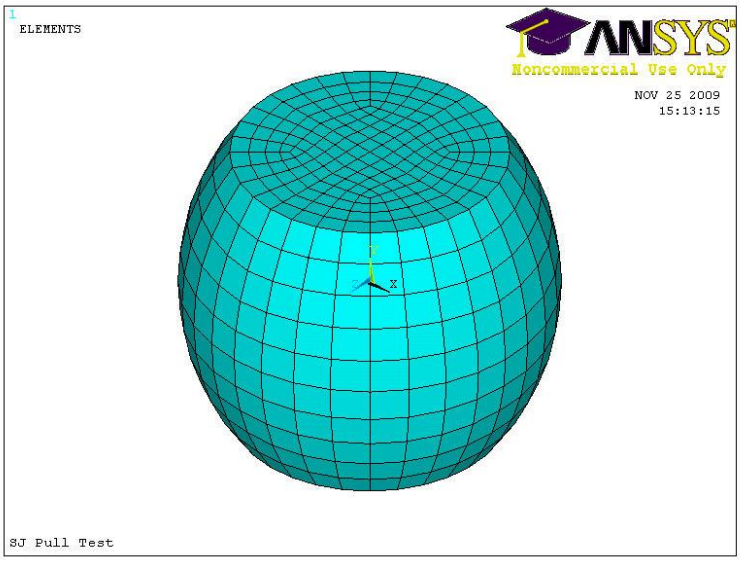


Figure 137: Meshed model of solder ball

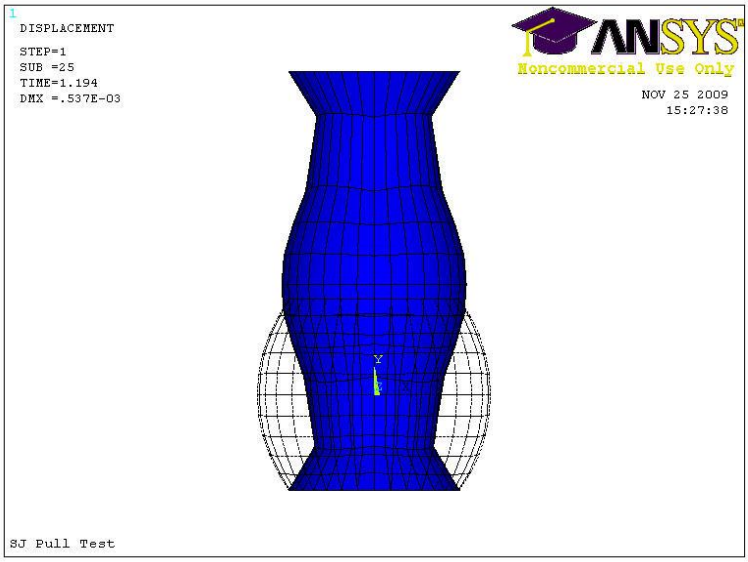


Figure 138: Deformed and undeformed geometry of solder ball

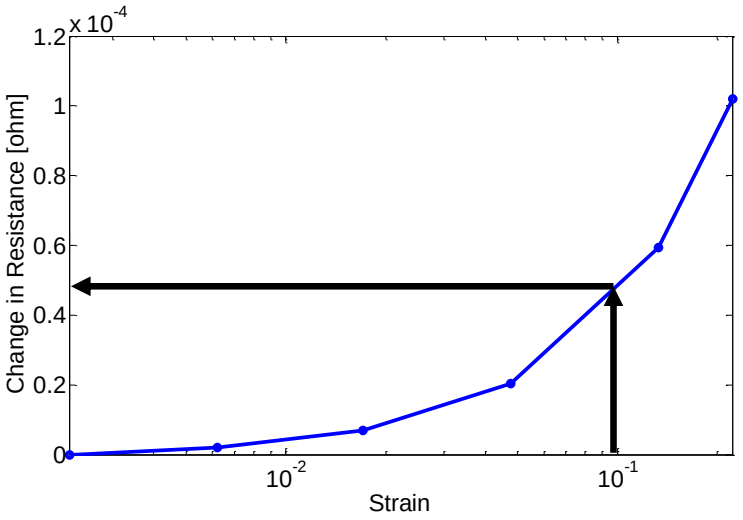


Figure 139: Simulated change in resistance of solder ball during pull test. Arrows indicate expected change in resistance at a strain of 0.1.

Resistance of the solder interconnect was computed by converting the VISOC107 elements to SOLID5 elements after intermediate steps in the deformation. A steady state conductance simulation was run using the deformed geometry after each sub-step. Using the built in macro command GMATRIX, the conductance of the solder ball in the deformed state could be calculated. The conductance is the inverse of the resistance. The meshed geometry before deformation can be seen in Figure 137, while the deformed geometry can be seen in Figure 138. Deformation was applied to the solder joint at a specified strain rate of 1 sec^{-1} typical of a shock test. An example of this mapping is shown in Figure 139.

Following a method similar to [Lall 2004c] the assumed criteria for failure in the simulated solder joint was based on the joint exceeding a critical plastic strain value. The critical plastic strain value was determined from a BGA pull test. Based on the experimental data at a strain rate of 1/sec, an overall strain of the solder joint of 0.1 corresponded to failure. Model predictions indicate a change in resistance of $5x10^{-5}$ Ω correlates with interconnect strain of 0.1 prior-to-failure of the interconnect. This critical resistance value derived from the FEM

simulation will be used as a threshold value to define failure for the PHM algorithm. Since the daisy chained resistance of a package is monitored in this study the critical resistance calculated from the FEM simulation must be scaled up from a single solder ball to account for changes in resistance of the entire package. This was achieved by approximating that every interconnect feels the same strain. Therefore the critical resistance is multiplied by the number of I/O in the package, i.e. 676 for the PBGA 676 to obtain the overall critical resistance value ($676x5x10^{-5}$ Ω = $3.38x10^{-2}$Ω). Assuming that every interconnect feels the same strain is not strictly correct since failure most often occurs in the corner interconnects. This implied averaging of strain across all interconnects is justified since deflections were small for the vibration test. Ultimately, errors from this approximation add uncertainty to the remaining useful life calculation. As will be demonstrated later in the paper, this uncertainty must be managed to obtain meaningful results.

9.4 Filtering and RUL Prediction

System damage state estimation in the presence of measurement noise and process noise has been achieved using the Kalman Filter. Previously, the Kalman Filter has been used in guidance and tracking applications [Kalman 1960, Zarchan 2000]. System state has been described in state space form using the measurement of the feature vector, the velocity of feature vector change and the acceleration of the feature vector change. System state at each future time has been computed based on the state space at the preceding time step, system dynamics matrix, control vector, control matrix, measurement matrix, measured vector, process noise and measurement noise. Figure 140 represents the data-flow through the system, where u_k is the control vector or input for the system, w_k is process noise, x_k is the state space vector at the k^{th} time step, H is the measurement matrix which is a constant in this implementation, v_k is the measurement noise, z_k the measured state, and Φ_k is the fundamental matrix.

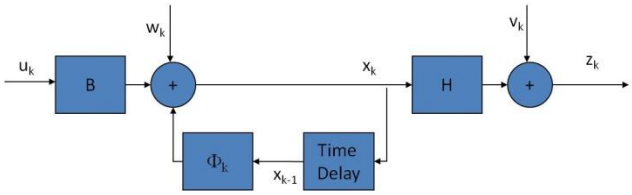


Figure 140:Graphical state space representation of a system

The equivalent Kalman Filter equation for state space representation in the presence of process noise and measurement noise is:

$$\bar{\mathrm{x}} = H\Phi_k\hat{x}_{k-1} - HG_k u_{k-1} \tag{238}$$

$$\hat{x}_k = \Phi_k\hat{x}_{k-1} + G_k u_{k-1} + K_k(z_k - \bar{\mathrm{x}}) \tag{239}$$

$$\mathrm{z_k} = Hx_k + v_k \tag{240}$$

where $\hat{x}_k$ is the Kalman Filter estimate of system-state at time k^{th} time step, x_k is the actual system state at the k^{th} time-step, and G_k is the control vector. The Kalman gain has been computed and updated at each time-step, while the filter is operating from the Riccati equations [Zarchan 2000]. The Ricatti equations can be represented in matrix form as:

$$M_k = \Phi_k P_{k-1}\Phi_k^T + Q_k \tag{241}$$

$$K_k = M_k H^T(HM_k H^T + R_k)^{-1} \tag{242}$$

$$P_k = (I - K_k H)M_k \tag{243}$$

where M_k is the covariance of errors in state estimates before update, Φ_k is the fundamental matrix which represents the system dynamics, Q_k is the discrete process noise matrix, K_k is the Kalman gain, H is the measurement matrix, and P_k is the covariance matrix representing errors in the state estimate after an update. R_k is the process noise matrix and has been used as a device for telling the filter that we know that filter's model of the real world is not precise. The diagonal elements of P_k represent variance of the true state minus the estimated

state. M_k is sometimes referred to as the a priori covariance matrix, and P_k may be referred to as the posterior covariance matrix.

The feature vector used for prognosis of the system health is not a constant or a straight line, therefore the zeroth and first order systems were ruled out and a second order system was used for representation of system state evolution with progression of underlying damage. The choice of the second order filter was also influenced by the general observation that feature vectors evolve non-linearly and generally accelerate towards the end of life. The rate of evolution of a second order system can be represented as follows:

$$\begin{Bmatrix} \dot{x} \\ \ddot{x} \\ \dddot{x} \end{Bmatrix} = [F] \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} x \\ \dot{x} \\ \ddot{x} \end{Bmatrix} \tag{244}$$

The fundamental matrix has been computed from the Taylor series expansion of the system dynamics matrix, F, as follows:

$$\Phi(t) = e^{Ft} = I + Ft + \frac{(Ft)^2}{2!} + \cdots + \frac{(Ft)^n}{n!} \tag{245}$$

$$\Phi(t) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} t + \frac{1}{2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} t^2 \tag{246}$$

$$= \begin{bmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}$$

$$\Phi_k = \Phi(T_s) = \begin{bmatrix} 1 & T_s & \frac{1}{2}{T_s}^2 \\ 0 & 1 & T_s \\ 0 & 0 & 1 \end{bmatrix} \tag{247}$$

Note that the F^3 terms and above are identically zero; therefore, the expansion only has three non-zero terms. A model based on the accrued plastic work in interconnects of the system has not been used because the inputs to the system are not always known or measurable and

cannot be assumed to always be constant or known in advance. Therefore the feature vector based on resistance spectroscopy has been related to the underlying plastic work and its evolution used for prognostication of system state and residual life. The first and second derivatives of the feature vector based on resistance spectroscopy have been computed to estimate the state of the feature vector at future time-steps. The system state vector is represented as $x_k = \lfloor x, \dot{x}, \ddot{x} \rfloor^T$, where x is the interconnect resistance of the daisy chained package, $\dot{x}$ is the ramp rate of the interconnect resistance, and $\ddot{x}$ is the second derivative with respect to time of the interconnect resistance. The state vector evolution is represented as follows:

$$\begin{Bmatrix} \dot{x}_{k+1} \\ \ddot{x}_{k+1} \\ \dddot{x}_{k+1} \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} x_k \\ \dot{x}_k \\ \ddot{x}_k \end{Bmatrix} \tag{248}$$

The uncertainty of each prediction was quantified using the posterior error covariance. As an engineering approximation the uncertainty is calculated using a straight line approximation. Then the uncertainty from the linear approximation is superimposed on the failure prediction obtained from iteratively solving for the intersection of a quadratic equation with the critical resistance threshold. This is a trade off in accuracy, for the benefit of algorithm simplicity.

Assuming that the feature vector and its first derivative are normal random variables (Gaussian), then a straight line approximation of the time to failure can be

$$t_f = \frac{x_f - \hat{x}(1)}{\hat{x}(2)} \tag{249}$$

where t_f is the time to failure, x_f is the failure threshold, $\hat{x}(1)$ is the estimated state of the system (resistance) and $\hat{x}(2)$ is the estimate of the first derivative. The numerator will have a variance equal to the variance of the position estimate, which is directly available in the posterior

error covariance matrix as P(1,1). The denominator will have a variance equal to the variance of the first derivative estimate, directly available as P(2,2). If

$$\sigma_F^2 = P(1,1) \tag{250}$$

$$\sigma_R^2 = P(2,2) \tag{251}$$

then it is demonstrated in [Swanson 2001] that the non-Gaussian distribution resulting from the ratio of two normal distributions with variances of σ_F^2 and σ_R^2 can be integrated to find the equivalent 68.4% probability range around the mean.

$$\sigma = 1.86\frac{\sigma_F}{\sigma_R} \tag{252}$$

The mean of the distribution from the straight line approximation is disregarded since the more iterative method was used to solve for the intersection of the quadratic equation with the critical resistance threshold.

The uncertainty estimate around the RUL prediction includes a number of simplifying assumptions about the nature of the system and should only be taken as a rough estimate. It will be shown later that this metric is still useful in understanding the operation of the algorithm and is necessary for fully utilizing the PHM framework to make risk based decisions.

The extrapolation of the estimated state into the future to determine the RUL was accomplished by using the state evolution equation to iteratively solve the intersection of a quadratic equation with the critical resistance threshold. The parameters of the quadratic equation are estimated from the Kalman filter.

The Kalman filter equations are recursive and must be initialized before the first measurement. The initial state estimate was taken as zero since the system is expected to have zero change in resistance before incurring damage. The measurement noise term was obtained

from the observed variance in the measurement system during the first 30 seconds of testing (R_k = 5e-6). The process noise term was taken as Q_k = 1e-9 and represents the uncertainty in the process dynamics. The diagonals of the posterior error covariance matrix were set arbitrarily large to 1000, which indicates a complete lack of trust in the initial state estimate. The filtering and prediction algorithm is summarized below.

Algorithm: Kalman Filtering and RUL prediction

1. Initialize variables at time step $t = 0$
2. Project state at the next time step, $\bar{x} = H\Phi_k\hat{x}_{k-1}$
3. Calculate error covariance before update, $M_k = \Phi_k P_{k-1}\Phi_k^T + Q_k$
4. Calculate Kalman gain, $K_k = M_k H^T (HM_k H^T + R_k)^{-1}$
5. Take measurement, $z_k = Hx_k$
6. Update estimate with measurement, $\hat{x}_k = \Phi_k\hat{x}_{k-1} + K_k(z_k - \bar{x})$
7. Calculate error covariance after measurement update, $P_k = (I - K_k H)M_k$
8. Extrapolate feature vector many time steps, n, to the failure threshold value, $\hat{x}_{k+n} = \Phi_{k+n-1}$
9. Report predicted RUL (and uncertainty)
10. Iterate to step 2 for next measurement (k = k +1)

9.5 Estimation of Remaining Useful Life

The Kalman filter tracking results used for prognostication are shown in Figure 141. The measured data has been obtained from resistance spectroscopy. The red line in the first plot is the state estimate from the Kalman filter. Note that the state estimate from the Kalman Filter is smoother than the raw data based feature vector. Smoothing facilitates faster convergence in the PHM algorithm. The lower two plots are estimates of the first and second derivative of the field quantity measured for construction of the feature vector. Any time the velocity is negative, the PHM algorithm cannot make a prediction. This causes the RUL predictions to oscillate before convergence. The convergence of the Kalman gain is shown in Figure 142.

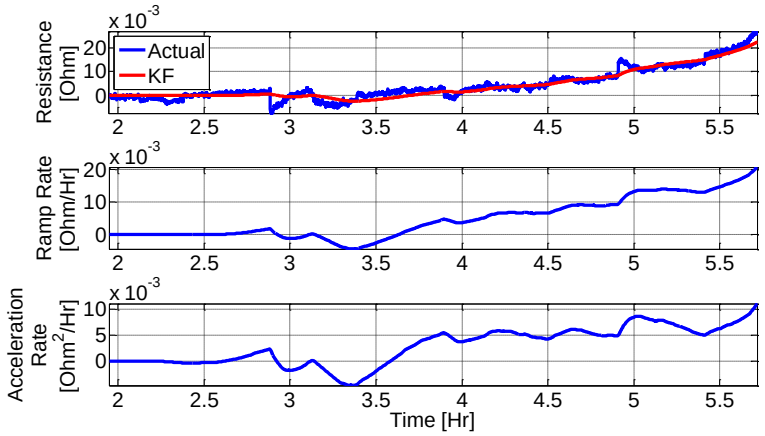


Figure 141: Results of Kalman Filtering

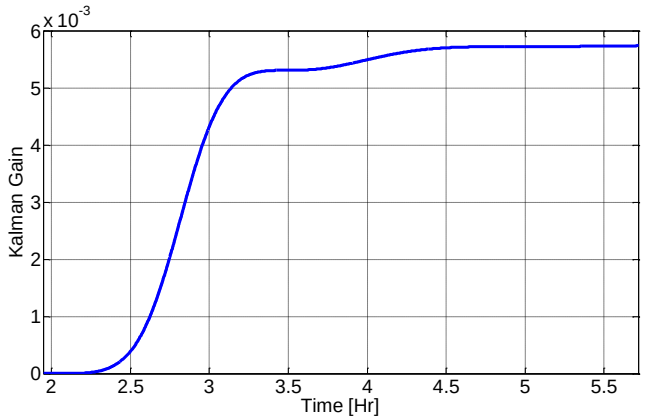


Figure 142: Convergence of the Kalman gain for the resistance estimate term.

The Newton-Raphson's method has been used for calculation of the RUL. A threshold value of $1x10^{-6}$ has been used as the threshold for convergence.

$$t_{n+1} = t_n - \frac{f(t)}{f'(t)} \tag{253}$$

$$t_{fn+1} = t_{fn} - \frac{\hat{x}(1) + \hat{x}(2)t_{fn} + 0.5\hat{x}(3)t_{fn}^2 - x_f}{\hat{x}(2) + \hat{x}(3)t_{fn}} \tag{254}$$

Where $f(t) = \hat{x}(1) + \hat{x}(2)t_{fn} + 0.5\hat{x}(3)t_{fn}^2 - x_f$, x is the state variable in the state space vector, t_{fn} is the estimate of the failure time at the time-step n, and x_f is the failure threshold for the state variable. The estimate of the failure time is updated in accordance with evolution of state-space vector which correlates with the underlying damage.

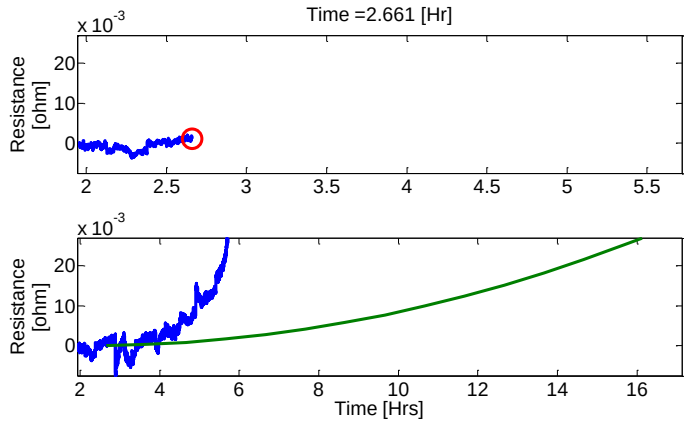


Figure 143: RUL prediction at 2.6 Hrs, the red circle shows what data was available for the prediction. The blue line in both plots is the feature vector, and the green line is the extrapolated state value used to predict RUL

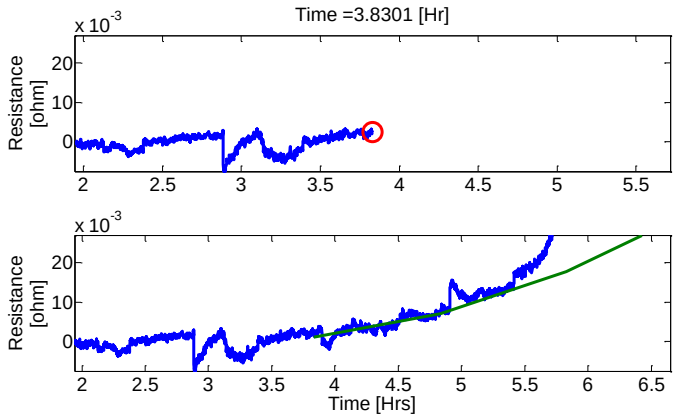


Figure 144: RUL prediction at 3.8 Hrs

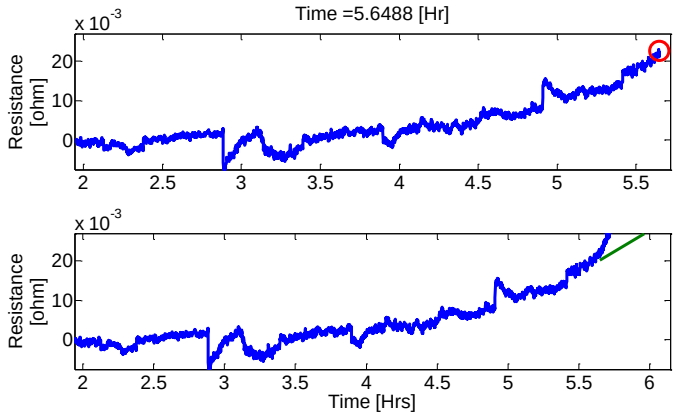


Figure 145: RUL prediction at 5.6 Hrs

The results of the RUL prediction are shown in Figure 143 Figure 144, and Figure 145.

Figure 143 is a prediction from early in the test. Based on the data available, which shows very

little change in the state variable resistance, the RUL prediction is considerably longer than the actual RUL. Figure 144 shows a prediction where more information is available to the algorithm. Figure 145, a prediction at the very end of the test shows the measured feature vector increasing in an exponential nature. The use of a quadratic model for predicting future states many time steps into the future was adequate for most of the test, but results in a prediction error at the extreme end of the test where failure propagation is highly non-linear. A higher order model or a non-linear model (implemented with an extended Kalman filter [Lall 2011a]) may provide better tracking and performance near the end of life, at the cost of implementation complexity. Using the PHM performance metrics, design decisions could be made to continue developing better system models, or to consider the error as an acceptable engineering approximation. A summary of all the RUL predictions compared against the actual RUL is shown in Figure 146. The initial estimates of the RUL oscillate and then gain traction in terms of accuracy following evolution of state space vector with underlying damage. PHM metrics for this implementation are shown in Figure 147 through Figure 149. The cost function for this implementation was 0.843.

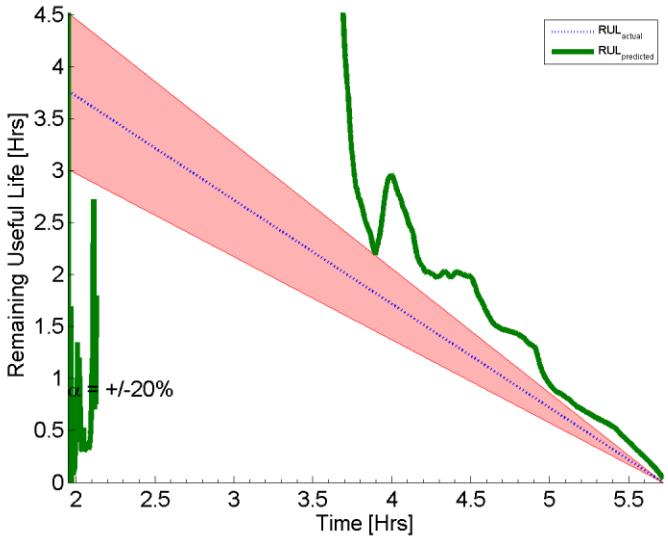


Figure 146: Comparison of actual RUL vs predicted RUL

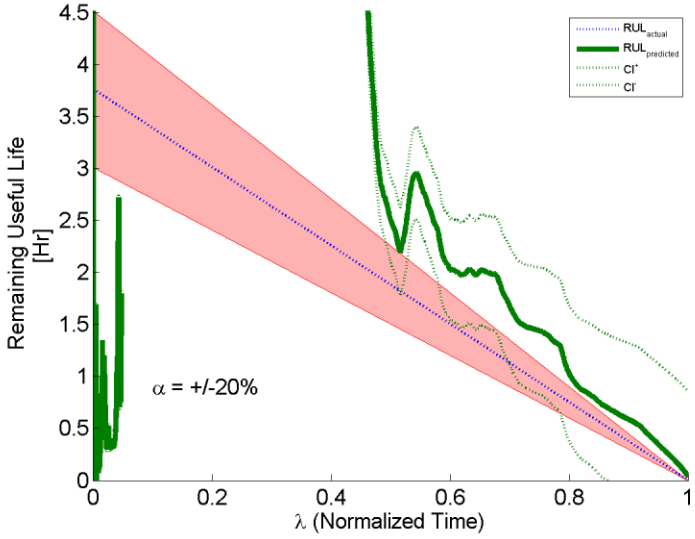


Figure 147: Alpha-Lambda accuracy metric illustrating the uncertainty surrounding RUL predictions

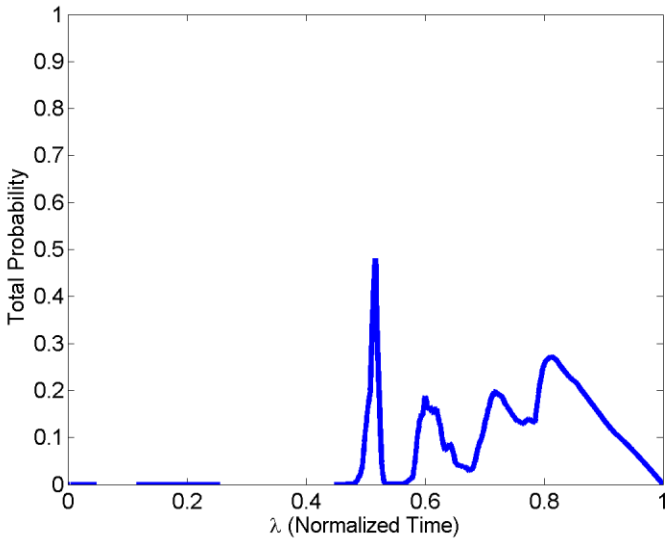


Figure 148: Beta metric illustrating the percentage of the RUL PDF that overlapped with the alpha bounds for each prediction

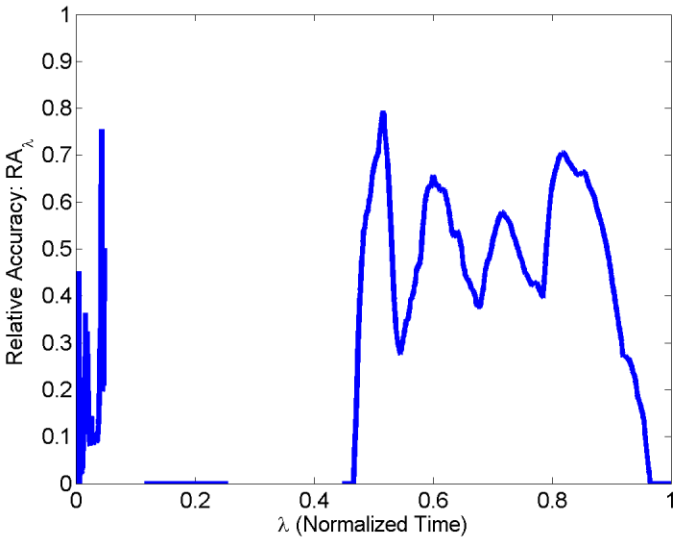


Figure 149: Relative Accuracy of the RUL predictions

9.6 Sensitivity Study

A sensitivity study was conducted to quantify the relationship each user definable parameter had on the performance of the PHM algorithm. Filtering methods are very sensitive to posterior "tweaking" that can improve reported performance. In practice a training data set would be required to enable prior knowledge of reasonable user definable parameters. To quantify the severity of incorrectly picking these parameters a cumulative beta was calculated to provide a single number that could represent the performance of the algorithm while user definable parameters were varied. Larger values of cumulative beta sum indicate better performance of the algorithm.

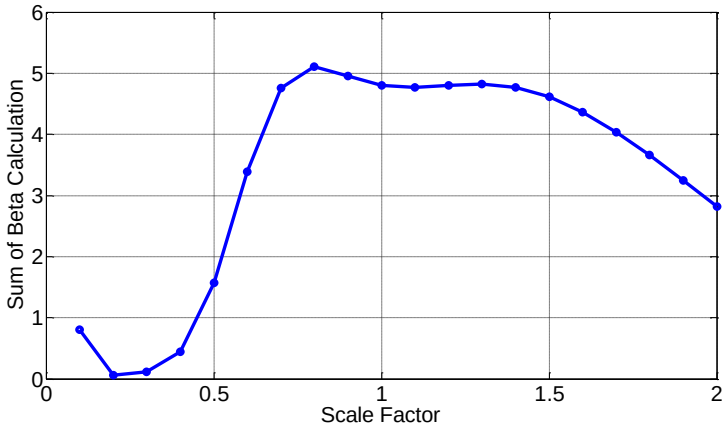


Figure 150: Variation in the sum of the beta calculation for variations in the critical-threshold of state variable.

Parameter variations studied include: state variable failure threshold and measurement noise in the Kalman filter. State variable failure threshold is the value at which the system is deemed to have failed. The process noise is a user-definable parameter signifying the underlying noise in the measured process. Increasing the measurement noise makes the estimated resistance measurement smoother, but less reactive to error between predicted and actual state values. Figure 150 shows the variation in the results of the algorithm when the critical value of the state variable was varied. Figure 151shows the cumulative beta with respect to the process noise.

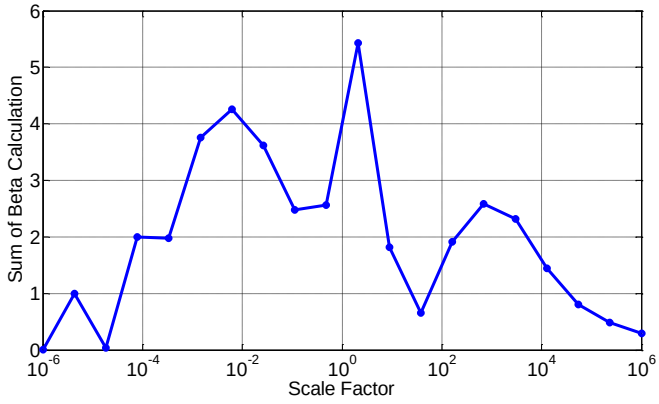


Figure 151: Variation in the sum of beta calculation for variations in tunable the process noise parameter

The sensitivity study shows that underestimating the critical value of state variable can severely hurt the performance of the PHM algorithm. A physics-based understanding of the degradation mechanism and its relationship to system performance is critical for implementation of the PHM algorithm. The cumulative beta score is less sensitive to process noise and therefore was varied over a number of orders of magnitude. An incorrect selection of either critical threshold for state variable or the process noise will have an adverse effect on the performance of the PHM algorithm.

9.7 Risk Based Decision Making

The practical result of predicting RUL is to make decisions. In the Bayesian framework used in this section, critical decisions about future use and replacement of a component can be justified using statistics. In an ultra-high reliability system, a critical decision is whether to replace a component. In high risk, mission critical systems for which this technique was designed, the maximum acceptable probability of failure is limited to 1%. This conservative restriction reflects the highly undesirable consequences of an unplanned failure.

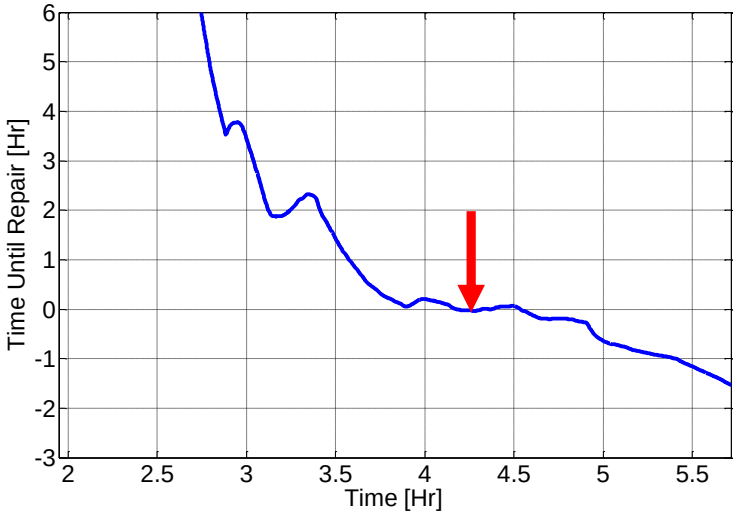


Figure 152: Time to order replacement component calculation vs time, which demonstrates one method of statistically defendable decision making using estimates of RUL.

The calculation to determine when to order a replacement part and schedule downtime for maintenance is based on the mean and standard deviation of the RUL prediction. In normally distributed data, the proportion of values within z standard deviations of the mean is

$$proportion = \text{erf}\left(\frac{z}{\sqrt{2}}\right) \tag{255}$$

where erf is the error function. A z value of 2.3263 represents the case where 98% of samples would be contained within +/-2.3263σ standard deviations of the mean. One percent of the samples outside +/-2.3263σ would fall on the negative side of the distribution, and the other one percent would fall on the positive side of the distribution. For predicting failures, we are only concerned with the negative side of the distribution or the one percent probability of failure. Using this approach, the appropriate time to order a replacement can be calculated. Assume that it takes 1-hour to order and receive a replacement component from the warehouse. Based on the predicted RUL, predicted RUL standard deviation, and a maximum acceptable probability of failure of one percent, the time until a replacement part should be reordered can be predicted by

$$t_{order} = RUL_{prediction} - 2.3263\sigma_{RUL} - t_{leadtime} \tag{256}$$

where σ_{RUL} is the standard deviation of the RUL, and $t_{leadtime}$ is the lead time for receiving the component after placement of the order. This equation is implemented on the data in this section and is shown in Figure 152. The order for the replacement component is placed when the t_{order} parameter reaches a value of zero, indicated by a red arrow in Figure 152.

9.8 Conclusion

A framework for prognosis of area-array electronics has been developed based on state-space vectors from resistance spectroscopy measurements, Kalman Filtering and Bayesian PHM framework. The measured state variable has been related to the underlying damage state by correlating the resistance change to the plastic strain accrued in interconnects using non-linear finite element analysis. The strain-resistance relationship has been used to define the critical resistance failure threshold for the component. The Kalman filter was used to estimate the state variable, rate of change of the state variable, acceleration of the state variable, and to construct a feature vector. The estimated state-space parameters were used to extrapolate the feature vector into the future and predict the time-to-failure at which the feature vector will cross the failure threshold. This procedure was repeated recursively until the component failed. Remaining useful life was calculated based on the evolution of the state space feature vector. Standard prognostic health management metrics were used to quantify the performance of the algorithm against the actual remaining useful life. An example application to part replacement decisions for ultra-high reliability systems was demonstrated. Finally the techniques described in the paper were used to determine the correct time to order a replacement for the component being monitored.

10 Extended Kalman Filter (EKF) Based Prognostic Algorithm Implemented for BGA's

A technique has been developed for monitoring the structural damage accrued in BGA interconnects during operation in vibration environments. The technique uses resistance spectroscopy based state space vectors, rate of change of the state variable, and acceleration of the state variable in conjunction with an Extended Kalman Filter and is intended for the pre-failure time-history of the component. Condition monitoring using the presented technique can provide knowledge of impending failure in high reliability applications where the risks associated with loss-of-functionality are too high to bear. The methodology has been demonstrated on SAC305 lead-free area-array electronic assemblies subjected to vibration. Future state of the system has been estimated based on a second order Extended Kalman Filter model and a Bayesian Framework. The measured state variable has been related to the underlying interconnect damage using plastic strain. Performance of the prognostication health management algorithm during the vibration test has been quantified using performance evaluation metrics. Model predictions have been correlated with experimental data. The presented approach is applicable to functional systems where corner interconnects in area-array packages may be often redundant. Prognostic metrics including alpha-lambda metric, beta, and relative accuracy have been used to assess the performance of the damage proxies. The presented approach enables the estimation of residual life based on level of risk averseness.

10.1 Introduction

Kalman filtering is a recursive algorithm that estimates the true state of a system based on noisy measurements [Kalman 1960, Zarchan 2000]. Prognostication of failure using Kalman

filtering has been demonstrated in steel bands and aircraft power generators [Batzel 2009, Swanson 2000, 2001]. Numerous applications in prognostics also exist for algorithms using more advanced filtering algorithms, known as particle filters. The state of charge of a battery was estimated and remaining useful life was predicted in [Saha 2009[a,b]]. Use of Extended Kalman Filtering for prognostication of electronic reliability based on the underlying damage mechanics is new. In this paper, a prognostic and health monitoring capability for electrical components based on changes in resistance has been presented. Failure modeling of BGA interconnects is combined with Extended Kalman filtering for plastic strain state estimation and a Bayesian framework for PHM. In contrast to the traditional Kalman filter, the extended Kalman filter allows the use of non-linear models to model damage in the interconnect being monitored. This work was an attempt to improve the results shown in the previous section by replacing the quadratic model with a more general exponential model. To make a fair comparison the data set used in the previous section will be used to hindcaste failure using the extended Kalman filter based PHM algorithm.

10.2 Test Vehicle

A set of test boards with multiple package architectures were used for experimental measurements. The test board includes package architectures such as plastic ball-grid arrays, chip-array ball-grid arrays, tape-array ball-grid arrays, and flex-substrate ball-grid arrays . The experimental matrix has ball counts in the range of 64 to 676 I/O, pitch sizes are in the range of 0.5mm to 1mm, and package sizes are in the range of 6mm to 27mm. The package parameters of this board are shown in Table 15. Representative sample of the test board is shown in Figure 153.

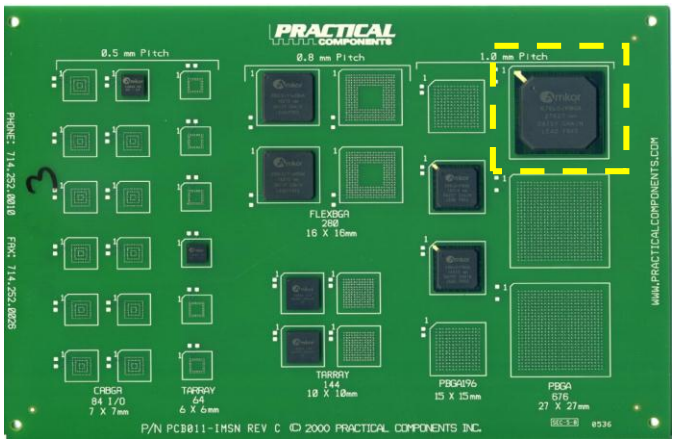


Figure 153: Test Board

Table 15: Package architectures on test board

	6 mm Tape Array	7 mm Chip Array	10 mm Tape array	15 mm PBGA	16 mm Flex BGA	27 mm PBGA
I/O	64	84	144	196	280	676
Pitch (mm)	0.5	0.5	0.8	1	0.8	1
Die Size (mm)	4	5.4	7	6.35	10	6.35
Substrate Thick (mm)	0.36	0.36	0.36	0.36	0.36	0.36
Pad Dia. (mm)	0.28	0.28	0.30	0.38	0.30	0.38
Substrate Pad	NSMD	NSMD	NSMD	SMD	NSMD	SMD
Ball Dia. (mm)	0.32	0.48	0.48	0.5	0.48	0.63

The test assemblies were subjected to vibration testing on a LDS Model V722 vibration table. A step stress profile was used to gradually ramp up the stress level to induce damage (Figure 132). The feature vector described in the previous section will be used for the

prognostics hindecaste. The same failure threshold of 3.125Ω obtained from the finite element simulation will again be used to define failure.

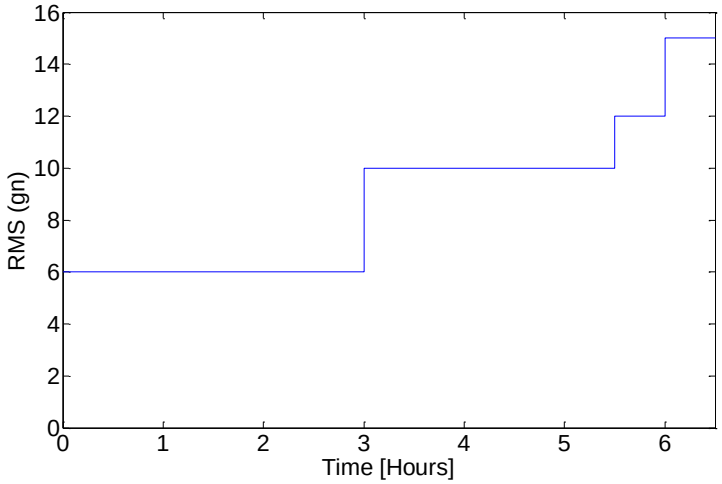


Figure 154: Step stress profile for vibration testing that fatigues interconnects to failure.

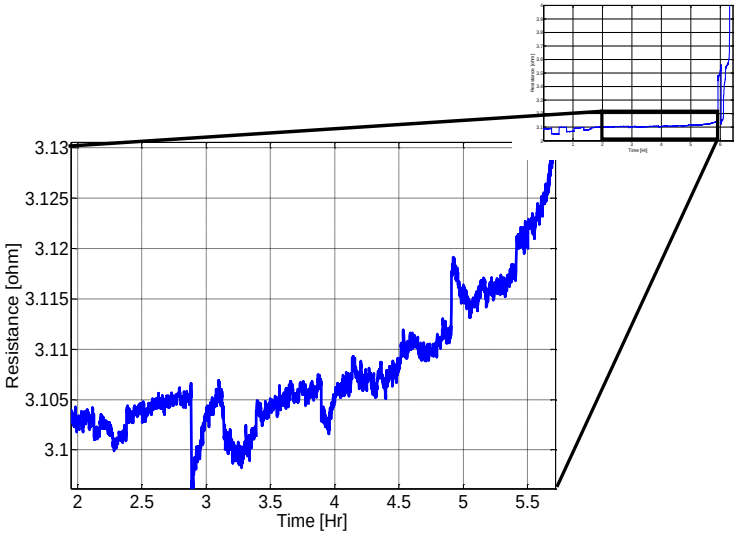


Figure 155: Zoomed view of resistance data between 2 hrs and failure

10.3 Filtering and RUL prediction

System damage state estimation in the presence of measurement noise and process noise has been achieved using the Extended Kalman Filter (EKF). System state has been described in state space form using the measurement of the feature vector, velocity of feature vector change and the acceleration of the feature vector change. System state at each future time has been computed based on the state space at preceding time step, system dynamics matrix, control vector, control matrix, measurement matrix, measured vector, process noise and measurement noise. The equivalent Extended Kalman Filter equation for state space representation is in the presence of process noise and measurement noise is:

$$\dot{x} = f(x) + w \tag{257}$$

Where x is the vector of system states being estimated, F is the system dynamics matrix that describes the evolution of the system, f(x) is a non-linear function of states, w is random zero mean noise defined as $w \sim N(0,\sigma_w)$. The system dynamics matrix, f, is a nonlinear function of the states. In extended Kalman filtering the relationship between system states (x) and measurements (z) can also be nonlinear, but in this paper are assumed to remain linear and occur at discrete time intervals, k.

$$z = h(x) + v \tag{258}$$

Where H is the measurement matrix, z is the measurement vector, h(x) is a measurement function which is a nonlinear function of states, v is zero-mean random process described by the measurement noise matrix. Since the system-dynamics (F) and measurement equations are nonlinear, a first-order approximation is used in the continuous Riccati equations for the systems dynamics matrix F and the measurement matrix H. The matrices are related to the nonlinear system and measurement equations according to

$$\mathrm{F} = \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x}} \tag{259}$$

$$\mathrm{H} = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}} \tag{260}$$

Where $\hat{x}$ is the Extended Kalman Filter estimate of system-state at the future time step. The Kalman gain has been computed and updated at each time-step, while the filter is operating from the Riccati equations [Zarchan 2000]. The Ricatti equations can be represented in matrix form as:

$$M_k = \Phi_k P_{k-1} \Phi_k^T + Q_k \tag{261}$$

$$K_k = M_k H^T (H M_k H^T + R_k)^{-1} \tag{262}$$

$$P_k = (I - K_k H) M_k \tag{263}$$

where M_k is the covariance of errors in state estimates before update, Φ_k is the fundamental matrix which represents the system dynamics, Q_k is the discrete process noise matrix, K_k is the Kalman gain, H is the measurement matrix, and P_k is the covariance matrix representing errors in the state estimate after an update. R_k is the process noise matrix and has been used as a device for telling the filter that we know that filter's model of the real world is not precise. The diagonal elements of P_k represent variance of the true state minus the estimated state. M_k is sometimes referred to as the a priori covariance matrix, and P_k may be referred to as the posterior covariance matrix.

The resistance of the package is measured directly in the experimental method. But the first and second derivatives are also desired to help extrapolate the state of the feature vector into the future. Simple numerical derivatives calculated from the raw feature vector are too noisy to be helpful. The Kalman filter is a powerful tool for smoothing and estimating the state of all of the desired variables. The general form of the resistance data is assumed to be.

$$x = ae^{bt} \tag{264}$$

Where a and b are constant parameters, t is time, e is Euler's constant and x is resistance. The derivative of resistance with respect to time and the b-parameter have been used to construct the state vector. The state vector is:

$$x_k = \lfloor x \quad \dot{x} \quad b \rfloor^T \tag{265}$$

Where, the resistance derivatives have been represented by:

$$\dot{x} = abe^{bt} = bx \tag{266}$$

$$\ddot{x} = ab^2e^{bt} = b^2x = b\dot{x} \tag{267}$$

$$\dot{b} = w \tag{268}$$

The state evolution equation is written as:

$$\begin{bmatrix} \Delta\dot{x} \\ \Delta\ddot{x} \\ \Delta\dot{b} \end{bmatrix} = \begin{bmatrix} \frac{\partial\dot{x}}{\partial x} & \frac{\partial\dot{x}}{\partial\dot{x}} & \frac{\partial\dot{x}}{\partial\ddot{x}} \\ \frac{\partial\ddot{x}}{\partial x} & \frac{\partial\ddot{x}}{\partial\dot{x}} & \frac{\partial\ddot{x}}{\partial\ddot{x}} \\ \frac{\partial\dot{b}}{\partial x} & \frac{\partial\dot{b}}{\partial\dot{x}} & \frac{\partial\dot{b}}{\partial\ddot{x}} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\dot{x} \\ \Delta b \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w \end{bmatrix} \tag{269}$$

Where w is a white process noise that has been added to the rate of change of acceleration equation for possible future protection. A model based on the accrued plastic strain in interconnects of the system has not been used because the inputs to the system are not always known or measurable and cannot be assumed to always be constant or known in advance. Therefore, the feature vector based on resistance spectroscopy has been related to the underlying plastic work and its evolution used for prognostication of system state and residual life. The derivatives of the feature vector based on resistance spectroscopy have been computed to estimate the state of the feature vector at future time-steps. The system dynamics matrix based on Equation (270) is given by:

$$\mathrm{F} = \left.\frac{\partial f(x)}{\partial x}\right|_{x=\hat{x}} = \begin{bmatrix} b & 1 & x \\ b^2 & b & 2bx \\ 0 & 0 & 0 \end{bmatrix} \tag{270}$$

The fundamental matrix has been computed from the Taylor series expansion of the system dynamics matrix, F, as follows:

$$\Phi(t) = e^{Ft} = I + Ft + \frac{(Ft)^2}{2!} + \cdots + \frac{(Ft)^n}{n!} \tag{271}$$

$$\begin{aligned} \Phi(t) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} b & 1 & x \\ b^2 & b & 2bx \\ 0 & 0 & 0 \end{bmatrix} t \\ &+ \begin{bmatrix} b & 1 & x \\ b^2 & b & 2bx \\ 0 & 0 & 0 \end{bmatrix}^T \begin{bmatrix} b & 1 & x \\ b^2 & b & 2bx \\ 0 & 0 & 0 \end{bmatrix} \frac{t^2}{2!} \end{aligned} \tag{272}$$

$$\Phi(T_s) = \begin{bmatrix} T_s^2\, b^2 + T_s b + 1 & bTs^2 + Ts & (3bxTs^2)/2 + xT_s \\ Ts^2 b^3 + T_s b^2 & Ts^2 b^2 + T_s b + 1 & (3xT_s^2 b^2)/2 + 2xT_s \\ 0 & 0 & 0 \end{bmatrix} \tag{273}$$

The discrete process noise matrix is described as:

$$\mathrm{Q_k} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Phi_s \end{bmatrix} \tag{274}$$

$$Q_k = \int_0^{T_s} \Phi(\tau) Q \Phi_k^T d\tau \tag{275}$$

As part of the Extended Kalman filter current state estimates are projected into the future. This was accomplished by Euler integration. In the following equations bars represent quantities projected into the future and hats represent current state estimates. All values with hats are estimated by the filter. The Euler integration time step, Ts, is the time step between measurements.

$$\bar{\ddot{x}} = \hat{x}(3)\hat{x}(2) \tag{276}$$

$$\overline{\dot{x}} = \hat{x}(2) + \overline{\ddot{x}} T_s \tag{277}$$

$$\overline{x} = \hat{x}(1) + \overline{\dot{x}} T_s \tag{278}$$

The filtering and prediction algorithm is summarized below.

Algorithm: Filtering and RUL prediction

11. Initialize variables at time step t = 0
12. Project state to the next time step, $\overline{x}_k$
13. Calculate error covariance before update, $M_k = \Phi_k P_{k-1} \Phi_k^T + Q_k$
14. Calculate Kalman gain, $M_k = \Phi_k P_{k-1} \Phi_k^T + Q_k$
15. Take measurement, z_k
16. Update estimate with measurement, $\hat{x}_k = \Phi_k \hat{x}_{k-1} + K_k(z_k - \bar{x})$
17. Calculate error covariance after measurement update, $P_k = (I - K_k H) M_k$
18. Extrapolate feature vector to failure threshold value using 'n' repeated applications of Euler integration, x_{k+n}
19. Report predicted RUL (and uncertainty)
20. Iterate to step 2 for next measurement (k = k +1)

10.4 Prognostic Hindecaste

To simulate the EKF based PHM algorithms ability to forecast failure, a prognostic hindcaste was performed using the data set described above. The results of Extended Kalman filtering are shown in Figure 156. The red line in the first plot is the state estimate from the Extended Kalman filter. Note that the state estimate from the Extended Kalman Filter is smoother than the raw data based feature vector. Smoothing facilitates faster convergence in the PHM algorithm. The lower two plots are estimates of the first derivative and the model parameter b which was introduced in equation (264) measured for construction of the feature vector. Any time the first derivative is estimated to be negative, the PHM algorithm cannot make a prediction and the algorithm reports a non-prediction error code. The middle trace in Figure 156 shows multiple occurrences where the derivative is estimated as negative. This causes the RUL predictions to oscillate before convergence. The convergence of the Kalman gain is shown in Figure 157.

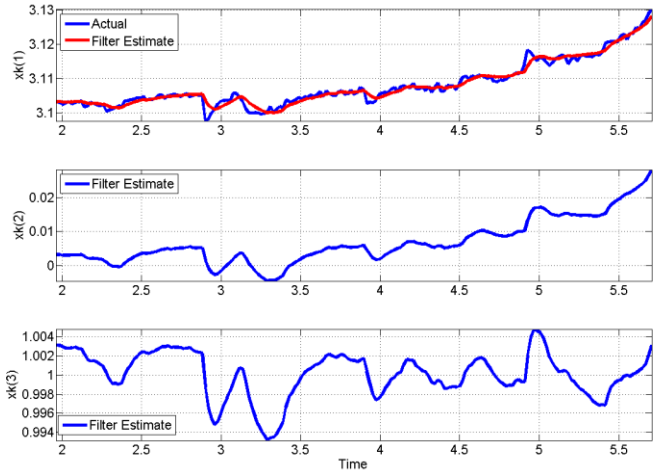


Figure 156: EKF filtering results

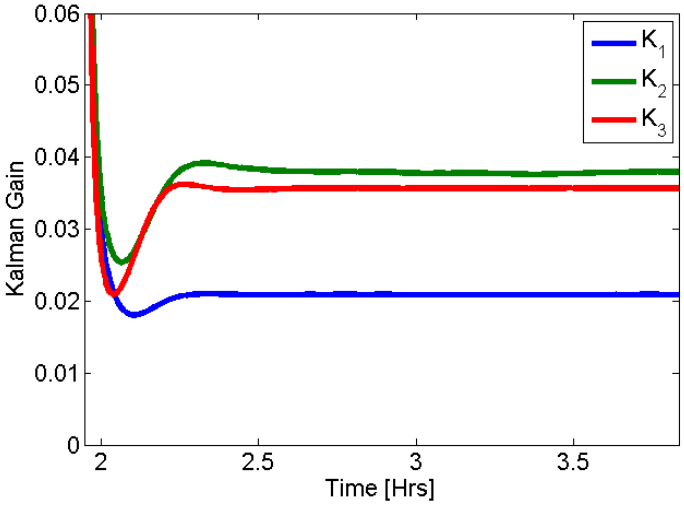


Figure 157: Convergence of Kalman gain

The Euler integration method has been used for calculation of the remaining useful life. Equations (276) through (278) are again used to project the current state estimate one time step into the future. If the projected state is greater than the failure threshold, the projected time step is taken as the time of failure. If the system state has not reached the failure threshold the following substitutions are made

$$\hat{x}(2) = \overline{\dot{x}} \tag{279}$$

$$\hat{x}(1) = \overline{x} \tag{280}$$

After substituting estimates back into the equation as state estimates the process in repeated ‘n’ times until the failure threshold is broken. The RUL is then calculated as the number of time steps required to propagate the system to failure multiplied by the length of the measurement time step. The calculation of RUL does not influence estimates provided by the filter. The uncertainty of each prediction was quantified using the posterior error covariance.

The results of the RUL prediction are shown in Figure 158, Figure 159, Figure 160. Figure 158 is a prediction from early in the test. Based on the data available, which shows very little change in the state variable resistance, the RUL prediction is considerably longer than the actual RUL. Figure 159 shows a prediction where more information is available to the algorithm. Figure 160, a prediction at the very end of the test shows how as the feature vector increases in an exponential nature that the assumption that the process dynamics are quadratic causes poor results. A summary of all the RUL predictions compared against the actual RUL is shown in Figure 161. The initial estimates of the remaining useful life oscillate and then gain traction in terms of accuracy following evolution of state space vector with underlying damage.

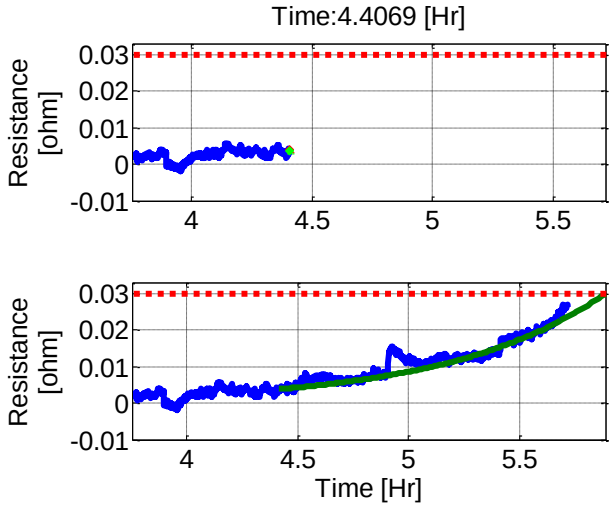


Figure 158: RUL prediction at 4.4069 Hrs, the red circle shows what data was available for the prediction. The blue line in both plots is the feature vector, and the green line is the extrapolated state value used to predict RUL

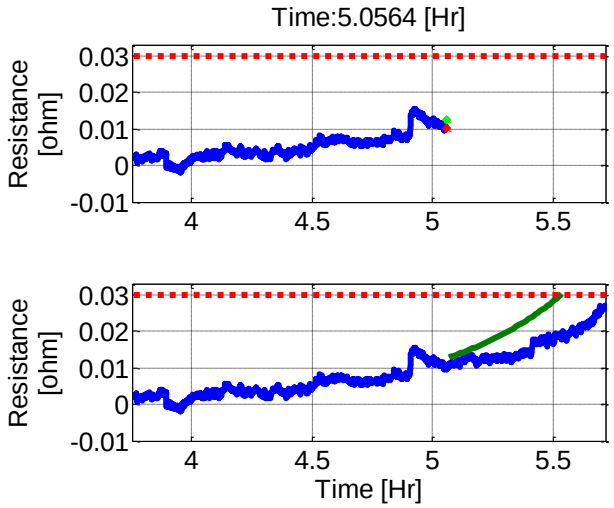


Figure 159: RUL prediction at 5.0564 Hrs

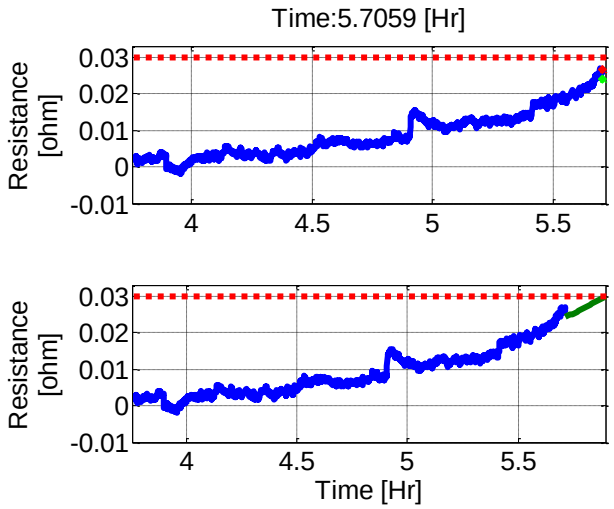


Figure 160: RUL prediction at 5.7059 Hrs

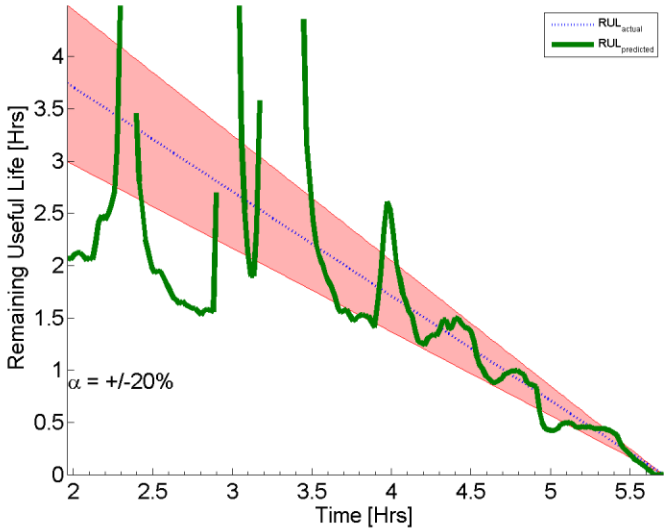


Figure 161: Comparison of actual RUL vs predicted RUL

Unreported RUL predictions around the time of 2.1,3 and 3.5 hours are the result of estimating the first derivative of the resistance as being negative. One of the advantages of this implementation is the ability to estimate model parameters in real time. The last state variable, denoted as model parameter 'b', was estimated online while the filter was recursively processing the data set. The residual between the one step ahead state estimate in the Kalman filter and the measurement is defined as

$$res_k = \bar{x}_k(1) - z_k \tag{281}$$

An error in the assumed model, or a sub optimal value for the model parameter can increase the residual. If the residual is positive the model is over predicting the dynamics of the system. If the residual is negative the model is under predicting the dynamics of the system. A plot of the residual compared to the estimated model parameter is shown in Figure 162.

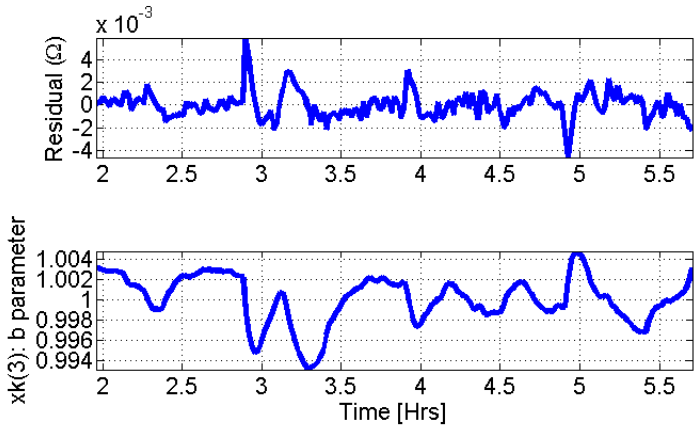


Figure 162: Residual between the estimated one step ahead prediction and the noisy measurement

It can be seen that when the residual is positive the filter decreases the estimate of the b parameter in an effort to match the observed true dynamics of the system. When the residual is negative the filter increases the estimate of the b parameter.

10.5 PHM Metrics

To facilitate a comparison of the presented algorithm against alternative implementation the standard PHM metrics were calculated for the EKF based PHM algorithm. The cost function for this implementation was 0.521.

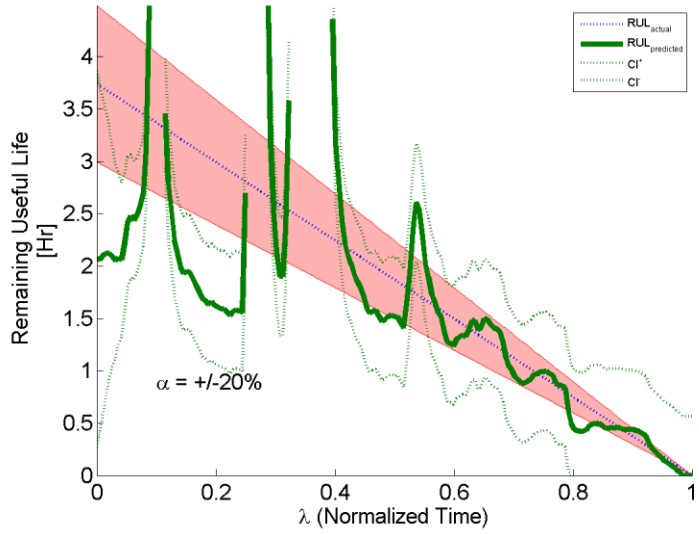


Figure 163: Alpha-Lambda Performance Metric

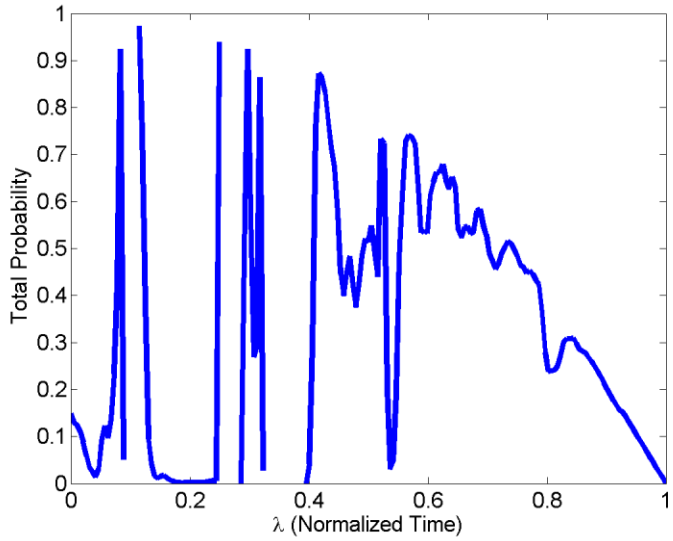


Figure 164: Beta Precision Metric

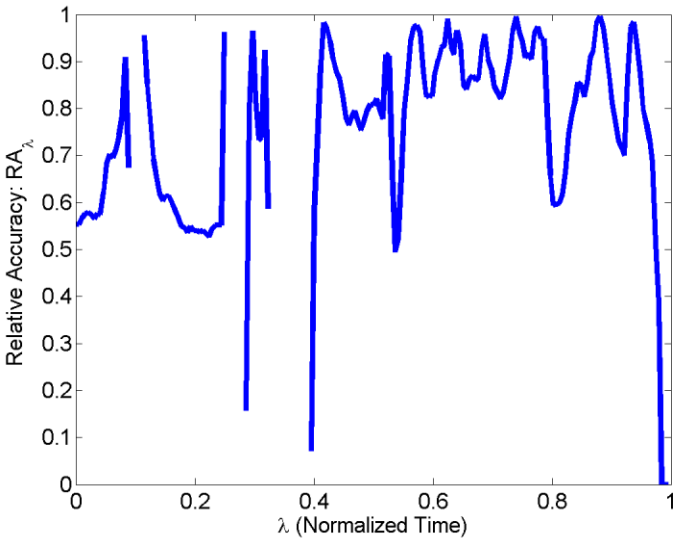


Figure 165: Relative Accuracy Metric

10.6 Comparison Against the Kalman Filter

To compare the performance of the Kalman filter based PHM algorithm and the EKF based algorithm the PHM metrics for each implementation have been reproduced in a side by side format. Besides the changes inherent with filtering with the Kalman and extended Kalman

filters, there were differences in the prognostic section of each algorithms. As discussed previously the underlying assumptions about the form of the equations generating the observed system were different. The Kalman filter used a closed form equation to predict RUL, while the EKF used Euler integration. As with linearizing the system model, there is also a risk of not using a small enough time step in the Euler integration step to propagate damage states forward. The uncertainty quantification approach was the same for both filters, but the reported values were not the same because different models were being used in the filters. The EKF is believed to a more flexible implementation because it has the ability to estimate online parameters in its own model. Lastly the EKF performed significantly better that the KF when judged by the PHM metrics. The cost function, which is used to summarize performance into a single number, shows that the EKF implementation was an improvement over the KF based PHM algorithm.

Component	Kalman Filter	Extended Kalman Filter
Model Order	Quadratic	Exponential
Prediction Mechanism	Closed form equation	Euler integration
Uncertainty Quantification	σ_F/σ_R	σ_F/σ_R
Parameter Estimation	NO	YES
Cost Function	0.843	0.521

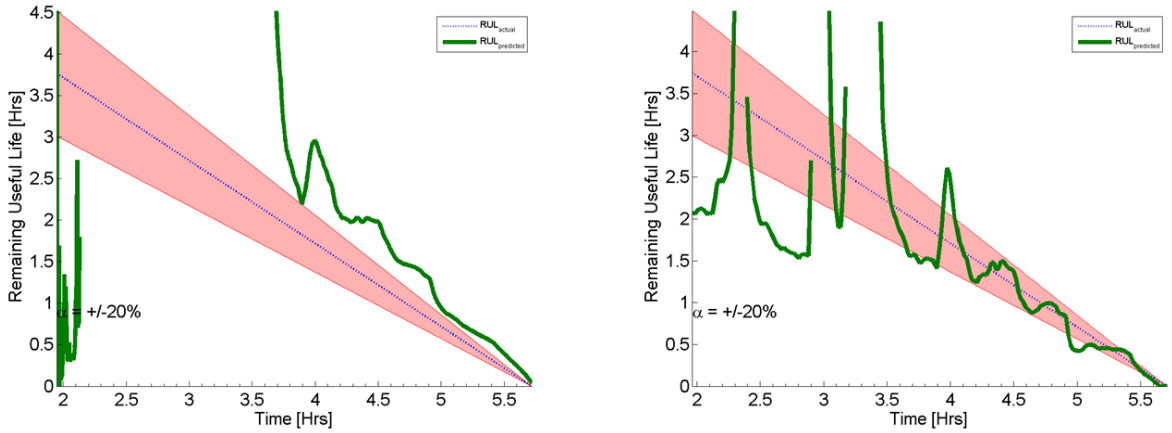


Figure 166: a) RUL plot for KF, b) RUL plot for EKF

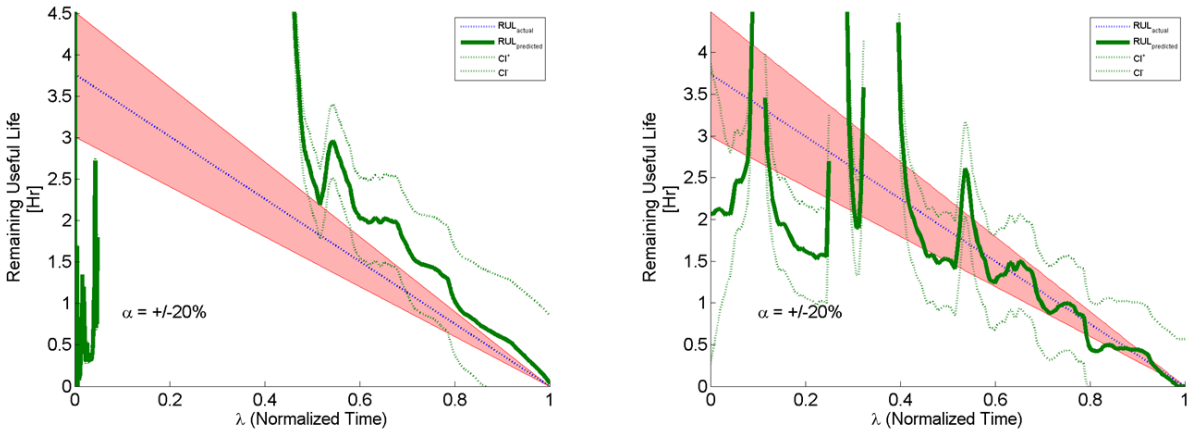


Figure 167: a) Alpha-Lambda plot for KF, b) Alpha-Lambda plot for EKF

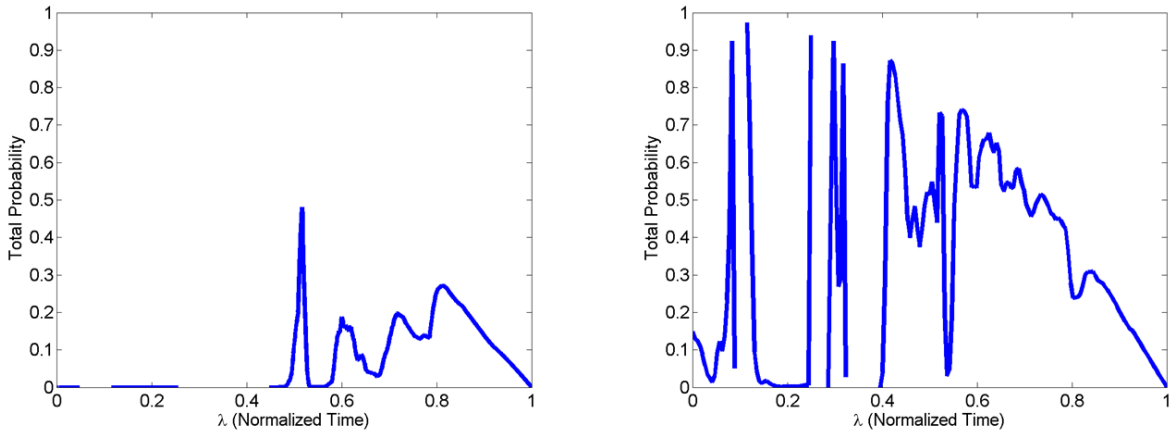


Figure 168: a) Beta metric for KF, b) beta metric for EKF

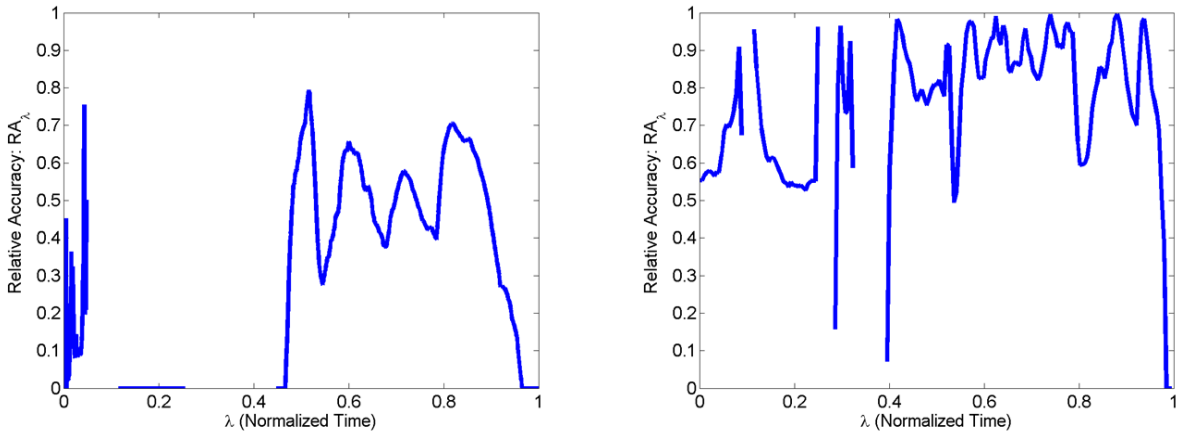


Figure 169: a) Relative accuracy for KF, b) Relative accuracy for EKF

11 Particle Filter (PF) Based Prognostic Algorithm Implemented for BGA's

In this section, the effectiveness of a proposed particle filter and resistance spectroscopy based approach in a prognostic health management framework has been demonstrated for electronics. The particle filter has been used to estimate the state variable, rate of change of the state variable, acceleration of the state variable and construct a feature vector. The estimated state-space parameters have been used to extrapolate the feature vector into the future and predict the time-to-failure at which the feature vector will cross the failure threshold. This procedure has been repeated recursively until the component failed. Remaining useful life has been calculated based on the evolution of the state space feature vector. Standard prognostic health management metrics were used to quantify the performance of the algorithm against the actual remaining useful life. Application to part replacement decisions for ultra-high reliability system has been demonstrated.

11.1 Introduction

Particle filtering is a recursive algorithm that estimates the true state of a system based on noisy measurements. Previously the particle filter has been used for projectile tracking [Arulampalam 2002, Gustafsson 2002, Ristic 2004], robot localization [Verma 2004], fatigue crack growth estimation [Cadini 2009], and prognostics for helicopters [Orchard 2007,2009], batteries [Saha 2009a,b], and pneumatic valves [Daigle 2009] . Particle filtering may also be known as sequential importance sampling, bootstrap filtering, or Monte Carlo filtering. In general particle filtering is classified as a sequential Monte Carlo technique [Doucet 2001]. The Kalman filter is a special case of the particle filter with constraints of linear Gaussian behavior on the underlying system and accompanying measurements. The benefits of using the particle

filter are a fully encompassed methodology for tracking non-linear/non-Gaussian damage variables with seamless incorporation of uncertainty management [Saha 2008] into the failure prediction algorithms.

11.2 Test Vehicle

A set of test boards with a single package architecture were used for experimental measurements. This unique assembly allows very precise failure analysis to be performed. The test board package architectures was a plastic ball-grid array with 324 I/O. The package parameters of this board are shown in Table 16. A representative sample of the test board is shown in Figure 170.

Figure 170: PBGA324-1mm-19mm test board

Table 16: Package Architectures used for Test Board B

Characteristic	Parameter
Package	Plastic BGA
Pitch	1mm
Body Size	19mm
Alloy	SAC305
Board surface finish	ImAg

Pad (Package)	SMD
Pad (Board)	NSMD
Daisy Chains	4
Total Daisy chain resistance	1.30Ω ± 0.2Ω
Package Manufacturer	Practical Components
Board Assembly	CAVE^3 , Auburn University

11.3 Test Environment

The test assemblies were mounted face down on an LDS LV217 electro-dynamic shaker table and subject to an 11gn random vibration profile, as shown in Figure 171. Continuity measurements for damage detection per IPC-SM785 and resistance spectroscopy measurements (both phase and magnitude) for prognostics were alternated every two seconds.

Testing was stopped when the traditional definition of failure which was taken as 10 intermittent events of >300Ω, each lasting >1μsec, per industry standard IPC-SM785 (1992 pg. 15-16) was met. Conveniently during testing resistance spectroscopy measurements alerted the experimenter that failure was imminent and to prepare to stop the test.

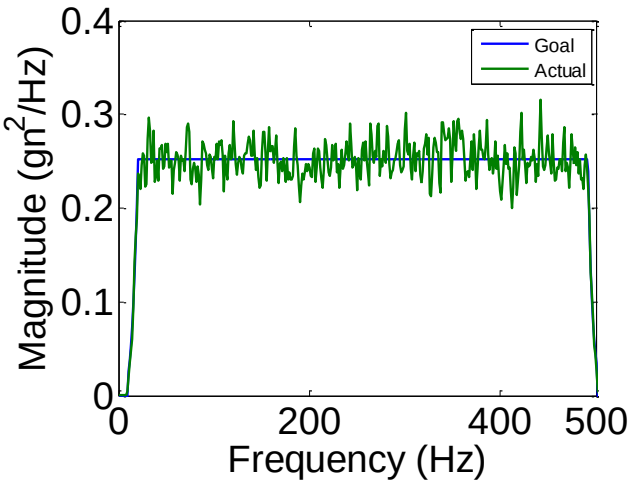


Figure 171: Vibration profile used for testing

11.4 Test Circuitry

The test circuitry used to monitor both resistance spectroscopy and daisy chain resistance comprised of two separate Wheatstone bridges that were switched between during testing using a single pole double throw relay. This setup effectively moves the device under test (daisy chained package) between the two measurement systems. The relay was transitioned between measurement devices every two seconds for the length of the test. The high level functional diagram of the test setup is shown in Figure 172.

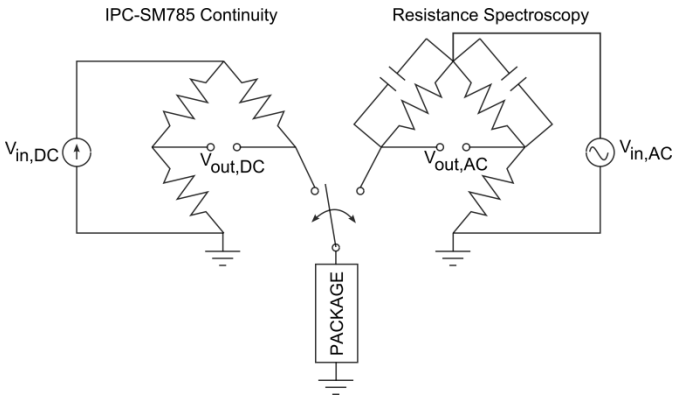


Figure 172: Functional diagram of test circuitry. The block labeled package represents the daisy chain resistance of the package under test

11.4.1 IPC-SM785 Continuity measurement

The electrical continuity was monitored using a LeCroy WR-Xi digital signal oscilloscope and a standard DC Wheatstone bridge. The daisy chained resistance of the package being tested was incorporated as an arm in the Wheatstone bridge. Changes in resistance of the package due to damage to solder interconnects is converted into a voltage through the bridge circuit at Vout,DC, and monitored with the digital signal oscilloscope. The sampling rate was set at 1x106 Samples/second to satisfy the requirement of high speed sampling. High resistance open events were automatically recorded to the internal hard drive of the oscilloscope for offline tallying. Every two seconds the continuity measurement was turned off to allow for a resistance spectroscopy measurement. The error introduced by this gap in measurement is considered acceptable since it only introduces a few seconds of uncertainty into the detected failure time for a test that lasts over an hour. It was observed that upon the initiation of failure, the occurrence of open events was very prevalent and nearly constant unlike thermal cycling failures which tend to be more intermittent in nature.

11.4.2 Resistance Spectroscopy

The experimental setup for a resistance spectroscopy measurement is similar to a continuity measurement, but utilizes additional equipment to detect very small changes in resistance that the continuity equipment does not have adequate resolution to detect. RS measurements are capable of detecting changes in resistance as small as a milli-ohm well before the traditional definition of failure and therefore contain prognostic value. A detailed diagram of the RS measurement setup is shown in Figure 4. Capacitors C1 and C2 help eliminate stray inductances from wires running between the test board and measurement equipment. Resistors R1, R2, and R3 are used to balance the bridge. The single pole double throw relay has a small

but non-negligible resistance that must be balanced out by specifying an appropriate value of resistor R3. Unlike traditional bridges, an AC voltage source drives this bridge, resulting in a sinusoidal output whose amplitude and phase shift are proportional to the resistance of the package. The lock-in amplifier performs the phase sensitive detection which effectively increases the resolution of the RS measurement compared to the continuity measurement. The outputs from the lock-in amplifier, the magnitude and phase shift of the signal $V_{out,AC}$, are recorded using a data logger.

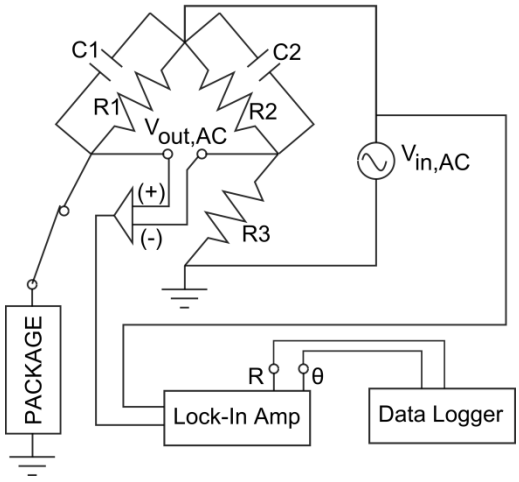


Figure 173: Resistance spectroscopy measurement setup. A differential output from the bridge is input into the lock-in amplifier for phase sensitive detection. Magnitude and phase data are recorded with the data logger.

Calculating the change in resistance of the package based on the measured output voltage from the bridge follows closely to that of a traditional Wheatstone bridge [Wheeler 2004], but must now incorporate impedances into the calculation. The general impedance bridge equation becomes

$$V_{out,AC} = V_{in,AC}\frac{Z_3Z_1 - Z_{PKG}Z_2}{(Z_2 + Z_3)(Z_1 + Z_4)} \tag{282}$$

Where Z_1 is the combined impedance of R1 and C1, and similarly for Z_2. Impedances Z_3 and Z_{PKG} are simply resistive impedances and therefore reduce to the value of resistor R3 and the

daisy chained resistance of the package respectively (i.e. Z_3=R3, Z_{PKG}=R_{PKG}). In a perfectly balanced bridge the numerator cancels out and the output voltage is zero. The resistance of the package can be solved for algebraically.

$$R_{PKG} = Z_{PKG} = \frac{V_{in,AC}Z_3Z_1 - V_{out,AC}(Z_2Z_1 + Z_3Z_1)}{V_{out,AC}(Z_2 + Z_3) + V_{in,AC}Z_2} \quad (283)$$

Table 17: Discrete component values used in resistance spectroscopy AC Wheatstone bridge

Component	Value
R_1,R_3	10Ω
R_2	1.28Ω
Initial R_{PKG}	1.30Ω
C_1,C_2	10 nF
$V_{in,AC}$	177.5 mV_{RMS} at 95 kHz

11.4.3 Phase Sensitive Detection

Inside the lock-in amplifier phase sensitive detection is used to measure very small changes in resistance which are converted to very small changes in voltage $V_{out,AC}$ by the bridge. Phase sensitive detection is a unique measurement method that allows the interrogation of very small signals corrupted with noise. Detection of signals with a signal to noise ratio considerably less than one is possible. Implementation involves multiplying a reference signal, $V_{in,AC}$, by the measured signal, $V_{out,AC}$. The resulting signal, V_m, is passed through a low pass filter to remove frequency components that are not of interest. In practical realizations the low pass filter is a higher order filter (e.g. 8^{th} order Butterworth) than depicted in the diagram. The output from the phase sensitive detection, V_{PSD}, is a signal that is proportional to the magnitude of the measured

signal, $V_{in,AC}$, without noise corruption. See [Lall 2009[b]] for a detailed discussion of phase sensitive detection including a derivation of pertinent equations. Since the components of the Wheatstone bridge are known quantities, the change in resistance of the package can be calculated to the milli-ohm resolution using the phase sensitive detection technique. A lock-in amplifier is a traditional method for performing the phase sensitive calculation, but other techniques based on switch type amplifiers or commercially available generic data acquisition hardware [Varcic 2001, Batra 2003] are alternative methods.

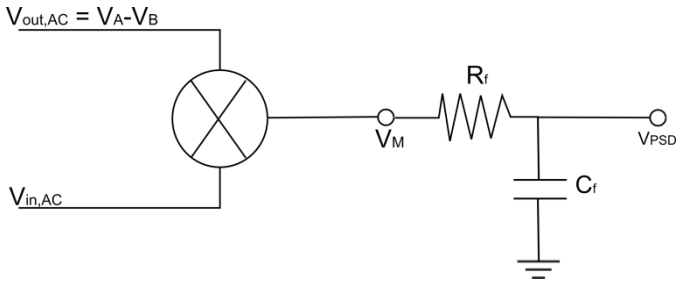


Figure 174: Phase sensitive detection performed inside the lock-in amplifier uses digital signal processing to multiply two sinusoids together and then eliminate high frequency components with a filter.

The resulting resistance measurement for the daisy chained package is shown in Figure 175. Notice how the resistance changes prior to the traditional definition of failure. This advance warning of failure validates the use of resistance spectroscopy as a prognostic indicator for failure in solder joint interconnects. If the test were allowed to continue past the point shown in the figure the resistance of the package would quickly rise to a very high level (>300Ω) permanently due to the violent nature of the test. In this test failure occurred at 70.55 minutes.

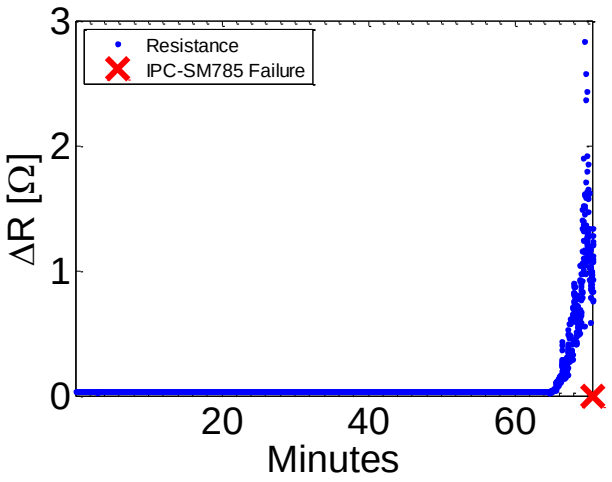


Figure 175: Change in resistance measured with the resistance spectroscopy method. Notice the significant change in resistance that is detectable well before the traditional definition of failure.

11.4.4 Transfer Function for Interconnect Strain to Resistance

To apply traditional strain based failure criteria [Lall 2007[f]] to solder interconnects the strain state in the interconnect is desired. Therefore a transfer function relating the measured resistance using the resistance spectroscopy technique to the non-observable strain state is needed. Previously [Lall 2009[a,b], 2010[a,b]] a correlation between interconnect strain and measured resistance was obtained using finite element modeling. This approach is advantageous for situations where no run to failure test data is available (e.g. new interconnect architectures, or projects still in the design phase). For this paper, the failure time is available from the IPC-SM785 continuity measurement will be used in place of simulation results. Previously, without IPC-SM785 continuity measurements, the actual resistance failure threshold has never been known as definitively.

11.5 Filtering and RUL Predictions

System damage state estimation in the presence of measurement noise and modeling uncertainty has been achieved using the sequential importance re-sampling (SIR) particle filter

[Arulampalam 2002]. In this implementation the monitored system state is the interconnect resistance which is a proxy for interconnect strain. The system state has been described in state space form using the measurement of the feature vector, velocity of feature vector change and the acceleration of the feature vector change. The estimated system state at each future time has been computed based on the state space at the preceding time step, and an assumed model for damage propagation. It is important to note that estimates and predictions are made in a recursive manner after each new measurement on the system and do not rely on prior knowledge of future measurements. Furthermore only one data measurement is stored in the system at a time and therefore requires minimal data storage. The preceding sub-sections will individually address damage modeling, damage tracking with the particle filter, and finally predicating remaining useful life.

11.5.1 Damage Modeling

In harsh environments, such as drop/shock and vibration, stress causes plastic deformation/strain to occur in solder interconnects. As described earlier traditional strain based failure criteria are used to describe failure in solder. Fatigue cracks initiate and propagate through local regions of sufficient plastic strain and are the observable result of plastic strain. From failure analysis it is known that the loss of electrical continuity in solder interconnects subjected to violent loadings is caused by a crack propagating through a multiphase material interface. Unfortunately this is a complicated process that is difficult to observe, which is why in this paper electrical resistance is used as a proxy. Experimental determination of fracture parameters has been demonstrated for relatively large geometries under quasi-static conditions [Marur 1999], but unfortunately solder material strength has been shown to be highly dependent on strain rate [Darveaux 2006]. Computationally expensive finite element methods have

correlated crack propagation with failure [Lall 2007[f]] but require a priori knowledge of failure locations. Extended finite element modeling (XFEM) techniques have been demonstrated for solder joint failure [Lall 2010[c]] that do not require explicit meshing of crack surfaces, but are still computationally expensive for real time applications. Material characterization for representative geometries and strain rates that are applicable to electronic packaging subjected to drop/shock and vibration are ongoing [Lall 2010[d]] and will someday be very useful for damage modeling. For the current work an empirical lumped parameter model based on exponential growth is used to model changes in resistance due to interconnect cracking. Traditionally a power law is used to model phase II crack growth. In this implementation phase III crack growth is also important since electrical continuity is not broken until crack propagation is completely across the interconnect severing electrical continuity. Therefore an exponential crack growth model [Mohanty 2009] capable of handling phase II and phase III crack growth is used.

Resistance of a cylindrical conductor is related to its respective geometry and material constants (ρ is electrical resistivity) through the relation

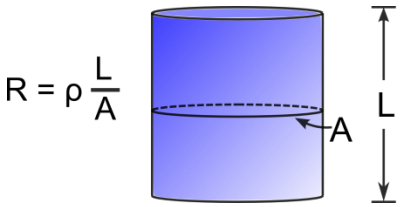


Figure 176: Theoretical relation between resistance, material properties, and geometry for a cylinder. A similar, more complicated relation exits for a solder interconnect

The relationship between resistance and geometry for a solder interconnect is more complicated and non-linear due to the shape of the interconnect but is expected to exist in a similar manner. As a crack propagates through the interconnect the effective area for conducting

current decreases in a non-linear fashion before the resistance ultimately spikes to infinity (a completely open crack).

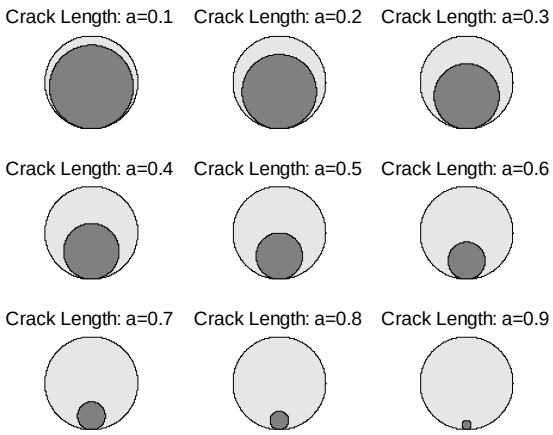


Figure 177: Proposed possible method of crack propagation in a solder interconnect resulting in a change in resistance. Electrical continuity is lost when crack length, a, reaches one.

Based on this intuition the resistance of the interconnect is assumed to grow exponentially, but the exact material and geometry effects are lumped into the model parameters. The state variable resistance, *x*(t), is modeled as

$$x(t) = ae^{bt} \tag{284}$$

$$\theta(t) = \begin{Bmatrix} a(t) \\ b(t) \end{Bmatrix} = \begin{Bmatrix} a \\ b \end{Bmatrix} \tag{285}$$

Where a and b combine together a number of material factors and geometry parameters into a single term. The model parameters can be a function of time, but are taken as constants to reduce complexity. Model parameters must be learned from existing run to failure data. The choice of model parameters is a decades old controversy inside the state estimation and filtering community which is discussed in the next section. The remaining state variables are simply derivatives of the damage model

$$\boldsymbol{x}(t) = \begin{Bmatrix} x(t) \\ \dot{x}(t) \\ \ddot{x}(t) \end{Bmatrix} = \begin{Bmatrix} ae^{bt} \\ abe^{bt} \\ ab^2e^{bt} \end{Bmatrix} = \begin{Bmatrix} x(t) \\ abx(t) \\ ab^2x(t) \end{Bmatrix} \tag{286}$$

11.5.2 Damage tracking

A SIR particle filter is used to track the damage state of the interconnect and provide probability distributions on state variables which are passed to the prediction algorithm. The particle filter represents the probability distribution over the current value of each state variable using a discrete probability mass function. Unlike with Kalman filtering there are no restrictions on the shape of the distribution function being approximated.

SIR Particle Filtering Algorithm

6. Initialize a distribution of particles, $\{x_0^i, w_0^i\}$
7. Project the particles forward one time step to find $p\left(x_{t_{p+1}} \middle| z_{0:t_p}\right)$
8. Make a new system measurement, z_{p+1}
9. Assign new weights based on the relevance of each particle compared to the new measurement $p\left(x_{t_{p+1}} \middle| z_{0:t_{p+1}}\right)$
10. Resample: If a few particles contain a majority of all possible weights, kill low probability particles and replace them with new more relevant particles

One difficult task when using any recursive filter is the choice of model parameters and noise terms. In the widely used Kalman filter this has been an acknowledged problem for at least 30 years [Grewal 2010], yet the Kalman filter has been applied successfully in a myriad of practical applications. The large number of possible fixes to the parameter problem with Kalman filters are cataloged in [Simon 2006]. The SIR particle filter suffers from the same problem. A possible solution is to include the model parameters into the state variable vector and allow the parameters to perform a random walk to identify appropriate values [Daigle 2009, Saha 2009[b], 2011a]. In this implementation fictitious process noise was added to account for ignorance about the other parameters. Table 18 summarizes the parameters used for the results shown in this paper.

Table 18: Parameter values used for damage model and particle filtering

Parameter	Value
a	0.5
b	0.4
σ_v	0.01
σ_R	$1x10^{-4}$
N	50

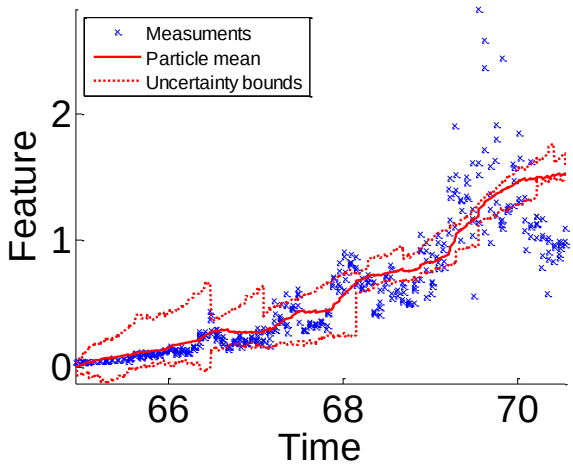


Figure 178: Results of particle filtering applied to feature vector shown in Figure 175

11.5.3 Damage Propagation and RUL prediction

The final remaining useful life prediction combines all of the algorithms described up to this point into a recursive algorithm that first calculate the predicted end of life (EOL) and then reports a remaining useful life (RUL) as a probability density. Using the latest state estimate from the particle filter each particle is projected forward in time until the failure threshold is crossed (C_{EOL}=1). The RUL prediction $p\left(RUL(t_p)\middle|z_{0:t_p}\right)$ is reported and the algorithm loops through to the next measurement. Damage states are propagated forward many time steps using Euler integration.

Damage Propagation and RUL Algorithm
for each Measurement

1. Initialize
   a. $\widetilde{x_k^i} = x_{t_p}^i$
   b. $\widetilde{\dot{x}_k^i} = \dot{x}_{t_p}^i$
   c. $\widetilde{\ddot{x}_k^i} = \ddot{x}_{t_p}^i$
2. While $C_{EOL} = 0$ for any $\tilde{x}_k^i$
   a. Propagate damage state forward
      i. $\widetilde{\ddot{x}_{k+1}^i} = ab^2\widetilde{x_k^i}$
      ii. $\widetilde{\dot{x}_{k+1}^i} = \widetilde{\dot{x}_k^i} + \Delta t\widetilde{\ddot{x}_k^i}$
      iii. $\widetilde{x_{k+1}^i} = \widetilde{x_k^i} + \Delta t\widetilde{\dot{x}_k^i} + \sigma_v$
      iv. $k = k + 1$
3. Report:
   $$p\left(EOL(t_p)\middle|z_{0:t_p}\right) \approx \min_{t \geq t_p}\{C_{EOL}(\tilde{x}_{k+t}^i) = 1\}$$
4. $RUL(t_p) = EOL(t_p) - t_p$

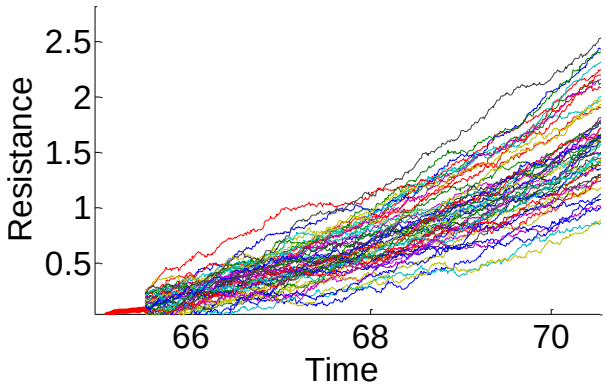


Figure 179: Damage propagation to predict EOL with a particle population size of N = 50

The complete algorithm that includes measuring damage proxies, tracking damage, and then predicting failures is repeated in a loop as long as it is deemed appropriate to continue running the system. A few of the predicted EOL distributions are shown in Figure 180.

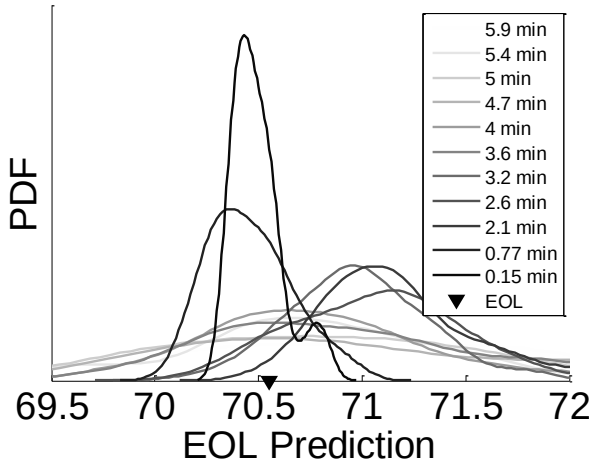


Figure 180: Predicted end of life distributions reported at decreasing times to failure. Darker lines represent predictions closer to the true end of life.

11.6 Prognostic Metrics

The resulting performance function for this implementation is 0.479.

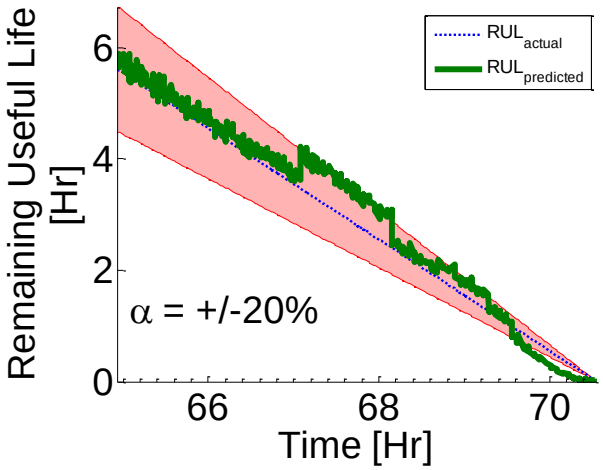


Figure 181: Remaining useful life predictions

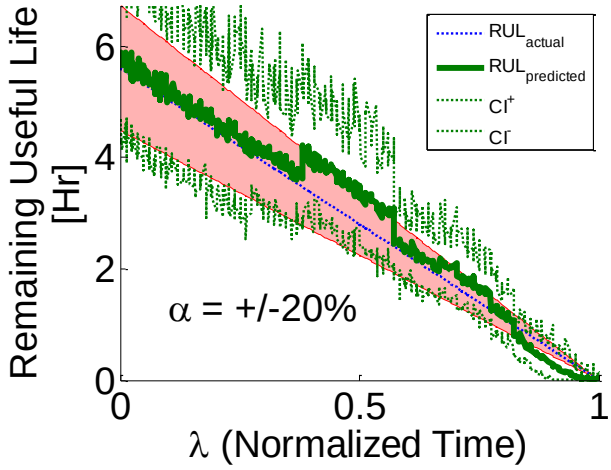


Figure 182: Alpha-Lambda Performance of PHM Algorithm

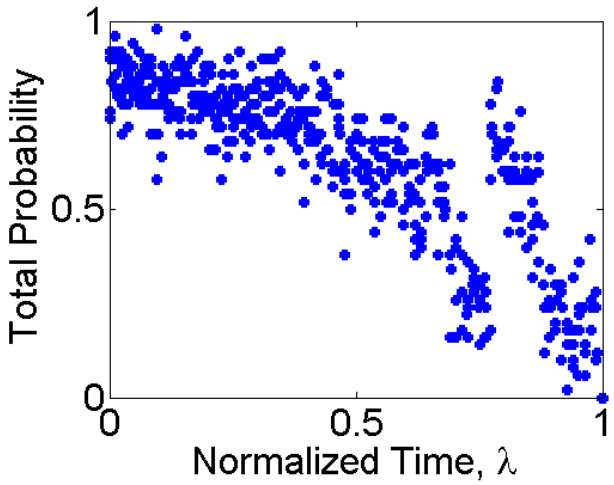


Figure 183: Beta calculation showing area under RUL prediction PDF that falls within the alpha bounds

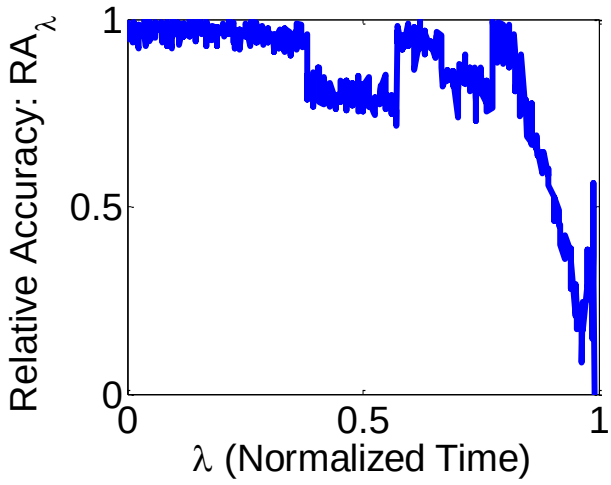


Figure 184: Relative Accuracy of RUL prediction

11.7 Comparison Against the Extended Kalman Filter

To compare the performance of the particle filter based PHM algorithm and the EKF based algorithm the PHM metrics for each implementation have been reproduced in a side by side format. The particle filter was implemented on the same data set used for the Kalman and Extended Kalman filter implementations. Unlike the changes inherent with filtering with the Kalman and extended Kalman filters, there were fewer differences in the prognostic section of the EKF and PF algorithms. The EKF linearizes the assumed exponential model, while the particle filter makes no linearizations or assumptions about Gaussian distributions. Without supplementing the particle filter with additional states or other ad-hoc methods, the ability of the particle filter to perform joint state parameter estimation is very limited. If the assumed model parameters are not close enough to actual values the particle filter can diverge and provide erroneous results. It is much more straight forward to perform joint state parameter estimation using the EKF. The uncertainty quantification approach was different for both filters. The EKF used the covariance matrix to quantify uncertainty, and the particle filter used a Monte Carlo implementation to propagate estimated states forward in time. The use of a Monte Carlo simulation seems like a better approach, but the impact that simplifying assumptions about uncertainty have on risk based decision making are hard to understand with the systems that are studied in this work.

Component	Particle Filter	Extended Kalman Filter
Model Order	Exponential	Exponential
Prediction Mechanism	Monte Carlo Euler integration	Single Point Euler Integration
Uncertainty Quantification	Based on Monte Carlo Results	σ_F/σ_R

Parameter Estimation	Limited	YES
Cost Function	0.479	0.521

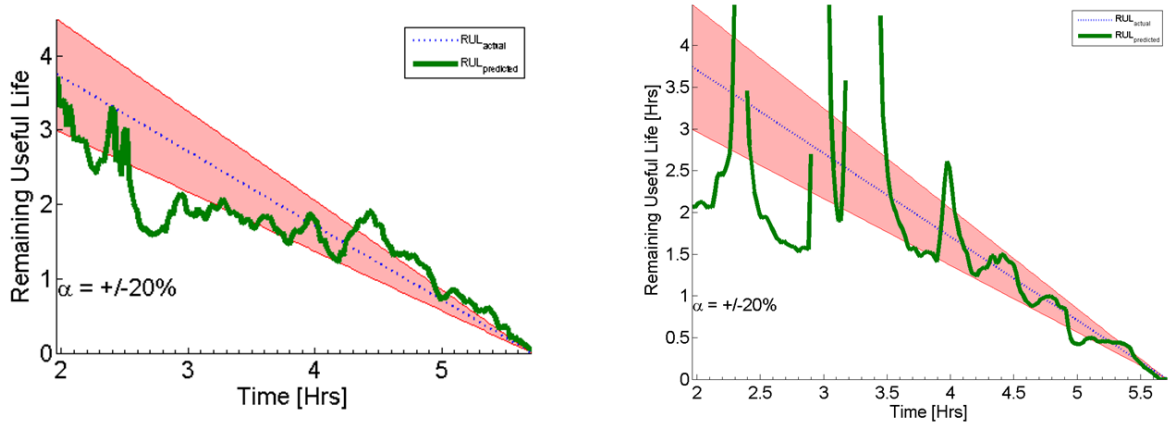


Figure 185: a) RUL plot for PF, b) RUL plot for EKF

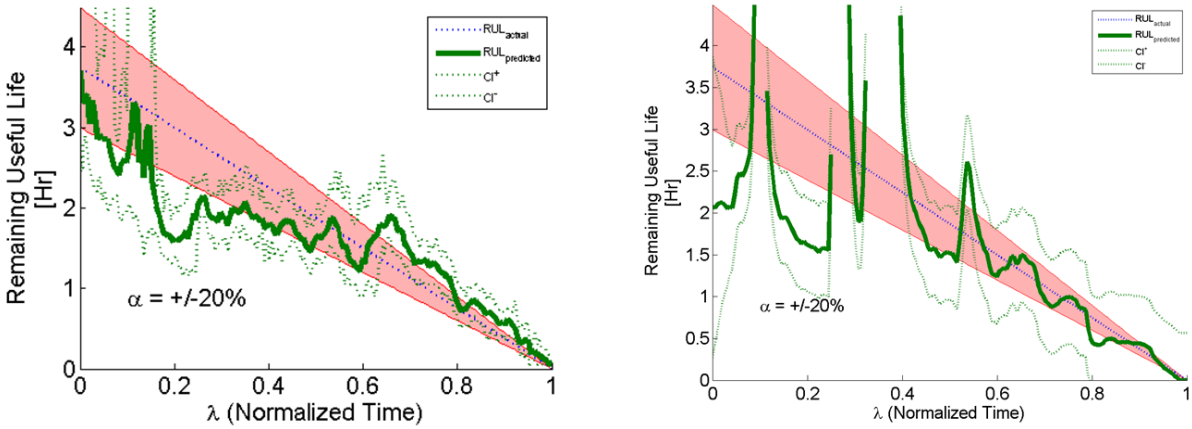


Figure 186: a) Alpha-Lambda plot for PF, b) Alpha-Lambda plot for EKF

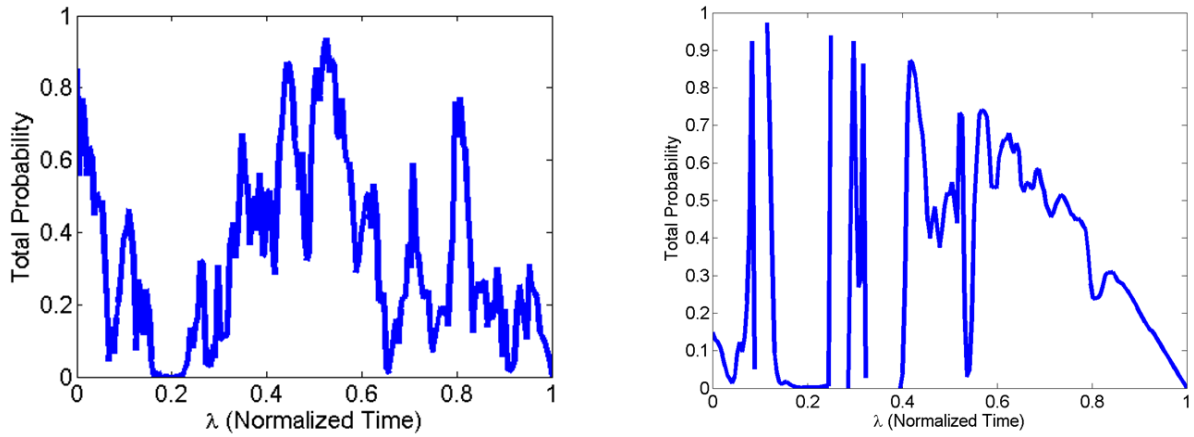


Figure 187: a) Beta metric for PF, b) beta metric for EKF

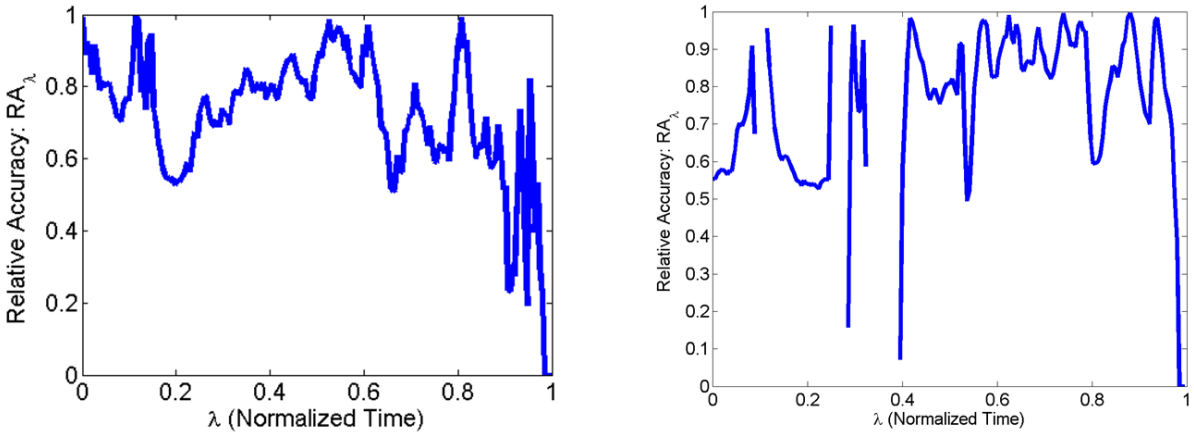


Figure 188: a) Relative accuracy for PF, b) Relative accuracy for EKF

Based on the PHM metrics the PF algorithm performs slightly better than the EKF implementation. A better implementation would be to combine the strengths of both algorithms. The EKF is better at online parameter estimation, while the uncertainty quantification using the particle filter is the superior approach.

11.8 Conclusion

In this paper a prognostic health management framework for electronics has been demonstrated that uses resistance spectroscopy measurements and a particle filter to track damage and predict remaining useful life. The test environment was a relatively violent vibration

test used for accelerated life testing of electronics. The prognostic value of resistance spectroscopy measurements has been quantified by monitoring for the traditional definition of failure (IPC-SM785) while simultaneously monitoring the solder interconnect using resistance spectroscopy. The use of the described algorithms in conjunction with described risk based decision making techniques provide a fully encompassed framework for tracking damage, managing uncertainty and making statistically defendable decisions. The outputs from the PHM algorithms presented here could be inputs into safety critical decisions about the future use of a product with an electrical subsystem. The algorithms could also support optimization of condition based maintenance for electronic rich systems. Standard performance evaluation metrics have been used to quantify the performance of the described PHM framework for electronics and has been shown to perform well compared to previous efforts.

12 PHM for electronics subjected to multiple simultaneous stress environments

Electronics in harsh environments may be subjected to extended periods of simultaneous high temperature and vibration. Further, electronics may be stored for prolonged periods of time prior to, and during deployment. Methods for assessment of accrued damage under simultaneous stresses are scarce. Test to failure data has been measured to study the effect of simultaneous thermal and vibration loadings on the reliability of BGA components. Two groups of pristine and isothermally aged components have been tested at both room temperature and 125°C while simultaneously being subjected to vibration loadings. The transient response of printed circuit boards under the overlapping stresses has been characterized. Damage accrued under overlapping stresses has been investigated using physics-based leading indicators of damage. The leading indicators are state vectors based on resistance spectroscopy and phase sensitive detection. An extended Kalman filter is employed to predict remaining useful life of the BGA components.

12.1 Introduction

Field deployed electronics, unlike controlled laboratory testing, may experience a variety of simultaneous mechanical stresses during its lifetime. In many applications such as automotive, aerospace, military, and electrical power generation, deployed electronics will experience simultaneous thermal and vibration stresses. By itself, single stress tests may not adequately characterize the reliability of an assembly. Further, usage may be preceded by prolonged storage in uncontrolled environments which can age assemblies. Traditional lifetime prediction methods may not accurately predict future performance if assemblies are deployed into unplanned usage

conditions that were not tested previously or if compounding reliability factors exist. This paper will present results from four test groups at a variety of test conditions. Lastly it is shown that resistance spectroscopy is a candidate for monitoring accrued damage in BGA components in simultaneous stress environments to provide real time reliability assurances. Bounds on the accuracy of prediction algorithms that could be expected in fielded components are presented.

Literature on simultaneous (aka: concurrent, overlapping) thermal and vibration stresses for BGA components is scarce compared to other standardized test methods such as temperature cycling, drop/shock, and vibration. There is no agreed upon test methodology, published results are contradictory, and direct comparisons are difficult to make. Nearly simultaneous testing was performed by [Rorgen 1998] for automotive electronics applications. [Zhao 2000] used moiré inferometry to conclude that solder joint strains were larger at elevated temperatures in cross sectioned SnPb assemblies subject to temperature and sinusoidal vibrations. The effect of thermal cycling and vibration on SAC305 LCCC, BGA, and resistors was studied by [Qi 2005]. Unfortunately the effect of temperature on the circuit board vibration response published in that paper could not be compared to [Matkowski 2005] because the boards and packages were significantly different. An interesting study on the variations in material properties with temperature of FR4 was given by [Hutapea 2003]. The authors argued that dampening from surrounding air was significant in calculating material properties so tests were performed inside a bell jar which was itself inside an oven. This highlights challenges in determining elevated temperature material properties for circuit boards. Also of note was the conclusion that the PCB shear modulus decreases by 50% between room temperature and 150°C. The elastic moduli decreases only slightly over the same temperature change.

While not directly applicable to simultaneous heating and vibration, literature concerning the effect of ageing on vibration only testing, and ageing/cycling on drop reliability provides insight into failure mechanisms and mechanics. Isothermal ageing decreases vibration only reliability [Tu 2001] and drop reliability [Ma 2007]. The effect of thermal cycling on drop reliability has been shown to increase reliability in some cases [Mattila 2006]. In [Mattila 2011] simultaneous heating and drop/shock loadings increased longitudinal strain on the PCB with increasing temperature. Further compounding the understanding of the mechanics and its effect on reliability is the effect of ageing [Ma 2009] and strain rate [Mattila 2005] in tensile tests on SAC solders.

Interrogation of system state and prognostication of accrued damage under simultaneous stresses is new, but previously for single stress conditions damage initiation, damage progression, and residual life in the pre-failure space has been correlated with micro-structural damage based proxies, feature vectors based on time, spectral and joint time-frequency characteristics of electronics [Lall2004a-d, 2005a-b, 2006a-f, 2007a-e, 2008a-f]. Methods applicable for the real time sequential processing of monitored damage indicators have been demonstrated for BGA interconnects subjected to single stress environments [Lall 2010[a-b], 2011[a-b]].

To characterize the transient response of the circuit board, digital image correlation has been used as a non-destructive method for monitoring in-situ, full field, circuit board strains during simultaneous temperature and vibration testing. Test vehicles include 324 I/O PBGA components assembled with SAC305 solder on two layer FR-4-06. Non-destructive, real time monitoring of solder joints was achieved with resistance spectroscopy [Lall 2009[a-c]]. Resistance

spectroscopy measurements were processed sequentially with an extended Kalman filter to facilitate the prediction of remaining useful life of components while being stressed.

12.2 Test Vehicle

A set of test boards with a single package architecture were used for experimental measurements. This unique assembly allows very precise failure analysis to be performed. The test board package architectures was a plastic ball-grid array with 19mm, 324 I/O, 1mm pitch package. The package parameters of this board are shown in

Table 19: Package Architectures used for Test Board

Characteristic	Parameter
Package	Plastic BGA
Pitch	1mm
Body Size	19mm
Alloy	SAC305
Board Surface Finish	ImAg
Pad (Package)	SMD
Pad (Board)	NSMD
Daisy Chain	4
Daisy Chain Resistance	1.3Ω±0.2Ω

Figure 189: PBGA324 test vehicle

12.3 Test Environment

The test assemblies were mounted face down on an LDS LV217 electro-dynamic shaker table surrounded by a thermal chamber and subject to an 11gn random vibration profile, as shown in Figure 2. Continuity measurements for damage detection per IPC-SM785 and resistance spectroscopy measurements (both phase and magnitude) for prognostics were alternated every two seconds. Testing was stopped when the traditional definition of failure which was taken as 10 intermittent events of >300Ω, each lasting >1μsec, per industry standard IPCSM785 (1992, pg. 15-16) was met. It was observed that once a package suffered its first intermittent open event, the remaining nine events occurred within a few seconds, and continued to occur until the input vibration was halted. Figure 192 highlights the effect of temperature on the mechanics of the printed circuit board. Non-contact deflection and strain measurements were obtained using digital image correlation [Lall 2007b] while simultaneous temperature and vibration testing was occurring. A window at the top of the thermal chamber provided *in-situ* optical access to the board.

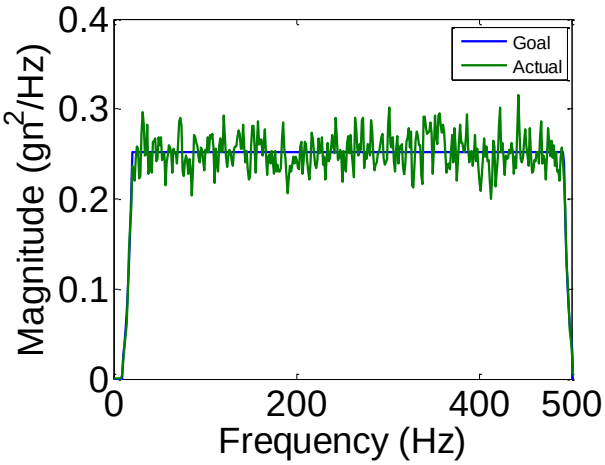


Figure 190: Vibration profile used for testing

Figure 191: Thermal chamber with upper window mounted on top of the shaker table

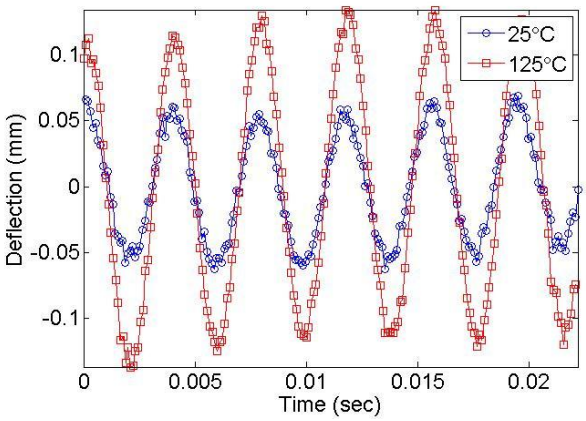


Figure 192: Deflection of the center of the board during vibration testing at different ambient temperatures

12.4 Test Circuitry

The test circuitry used to monitor both resistance spectroscopy and daisy chain resistance comprised of two separate Wheatstone bridges that were switched between during testing using a single pole double throw relay. This setup effectively moves the device under test (daisy chained package) between the two measurement systems. The relay was transitioned between measurement devices every two seconds for the length of the test. The high level functional diagram of the test setup is shown in Figure 193.

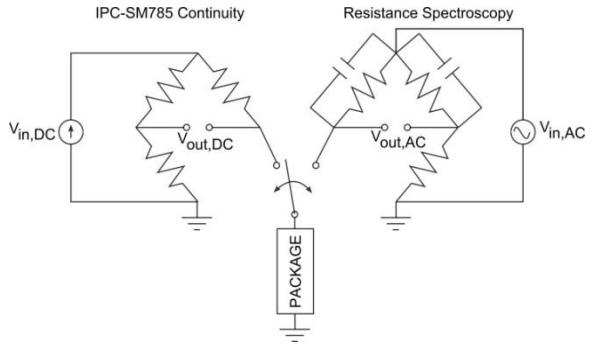


Figure 193: Functional diagram of test circuitry. The block labeled package represents the daisy chain resistance of the package under test

12.4.1 IPC-SM785 Continuity measurement

The electrical continuity was monitored using a LeCroy WR-Xi digital signal oscilloscope and a standard DC Wheatstone bridge. The daisy chained resistance of the package being tested was incorporated as an arm in the Wheatstone bridge. Changes in resistance of the package due to damage to solder interconnects is converted into a voltage through the bridge circuit at $V_{out,DC}$, and monitored with the digital signal oscilloscope. The sampling rate was set at $1x10^6$ Samples/second to satisfy the requirement of high speed sampling. High resistance open events were automatically recorded to the digital storage oscilloscope for offline tallying. Continuity measurement and resistance spectroscopy measurement were alternatively made every two-seconds. It was observed that upon the initiation of failure, the occurrence of open

events was very prevalent and nearly constant unlike thermal cycling failures which tend to be more intermittent in nature.

12.4.2 Resistance Spectroscopy

The experimental setup for a resistance spectroscopy (RS) measurement is similar to a continuity measurement, but utilizes additional equipment to detect very small changes in resistance that the continuity equipment does not have adequate resolution to detect. RS measurements are capable of detecting changes in resistance as small as a milli-ohm well before the traditional definition of failure and therefore contain prognostic value. A detailed diagram of the RS measurement setup is shown in Figure 4. Capacitors C_1 and C_2 help eliminate stray inductances from wires running between the test board and measurement equipment. Resistors R_1, R_2, and R_3 are used to balance the bridge. The single pole double throw relay has a small but non-negligible resistance that must be balanced out by specifying an appropriate value of resistor R3. Unlike traditional bridges, an AC voltage source drives this bridge, resulting in a sinusoidal output whose amplitude and phase shift are proportional to the resistance of the package. The lock-in amplifier performs the phase sensitive detection which effectively increases the resolution of the RS measurement compared to the continuity measurement. The outputs from the lock-in amplifier, the magnitude and phase shift of the signal $V_{out,AC}$, are recorded using a data logger.

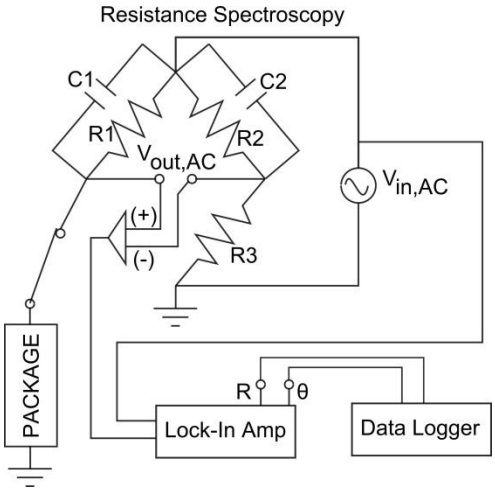


Figure 194: Resistance spectroscopy measurement setup. A differential output from the bridge is input into the lock-in amplifier for phase sensitive detection. Magnitude and phase data are recorded with the data logger.

Component	Value
R1,R3	10Ω
R2	1.28Ω
Initial Rpkg	1.30Ω
C1,C2	10 nF
Vin,AC	177.5mVRMS at 95 kHz

Figure 195: Discrete component values used in resistance spectroscopy AC Wheatstone bridge

12.4.3 Phase Sensitive Detection

Phase sensitive detection has been used to measure very small changes in resistance which are converted to very small changes in voltage $V_{out,AC}$ by the bridge. Phase sensitive detection is a unique measurement method that allows the interrogation of very small signals corrupted with noise. Detection of signals with a signal to noise ratio considerably less than one is possible. Implementation involves multiplying a reference signal, $V_{in,AC}$, by the measured

signal, $V_{out,AC}$. The resulting signal, V_m, is passed through a low pass filter to remove frequency components that are not of interest. In practical realizations the low pass filter is a higher order filter (e.g. 8th order Butterworth) than depicted in the diagram. The output from the phase sensitive detection, V_{PSD}, is a signal that is proportional to the magnitude of the measured signal, $V_{in,AC}$, without noise corruption. See [Lall 2009b] for a detailed discussion of phase sensitive detection including a derivation of pertinent equations. Since the components of the Wheatstone bridge are known quantities, the change in resistance of the package can be calculated to the milli-ohm resolution using the phase sensitive detection technique. A lock-in amplifier is a traditional method for performing the phase sensitive calculation, but other techniques based on switch type amplifiers or commercially available generic data acquisition hardware [Varcic 2001, Batra 2003] are alternative methods.

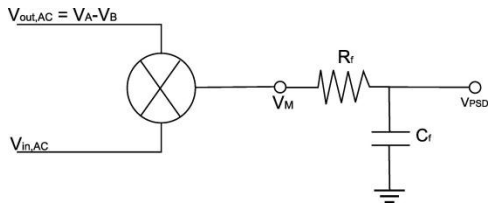


Figure 196: Phase sensitive detection performed inside the lock-in amplifier uses digital signal processing to multiply two sinusoids together and then eliminate high frequency components with a filter

12.5 Simultaneous Temperature and Vibration Testing

A 2x2 full factorial test matrix was implemented to study the effect of simultaneous temperature and vibration loadings on the reliability of BGA components. Vibration loadings were held constant as an 11gn random vibration profile in all tests. Ambient temperatures were chosen at two levels, 25C and 125C. Two levels of ageing conditions were studied, pristine (no ageing) and isothermally aged at 125C for 7 days.

Table 20: Full factorial test matrix used to study effect of simultaneous stresses and ageing effects

	Pristine	Aged (125C for 7 days)
11gn random vibration at 25C	Pristine-25C	Aged-25C
11gn random vibration at 125C	Pristine-125C	Aged-125C

Test vehicles were drawn from the same lot of PCB's and assembled components. Test conditions were alternated between 25C and 125C to help negate the effects of systematic errors in the experimental procedure. Time histories of the component daisy chain resistance measured using resistance spectroscopy are shown in Figure 197through Figure 200. IPC failure definitions are overlaid on the time histories. Each group consisted of five individual tests, but all five time histories may not be clearly visible in each plot. A summary of failure times and corresponding Weibull statistics is shown in Table 21. Weibull statistics were obtained by fitting models to experimental data using a maximum likelihood parameter estimation approach.

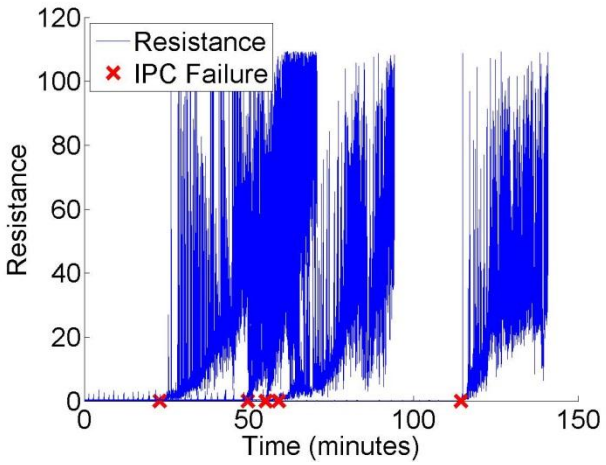


Figure 197: Time history of daisy chain resistance for pristine boards tested at 25C

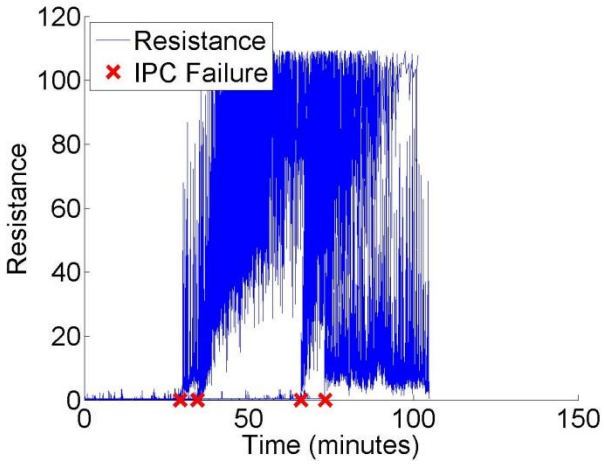


Figure 198: Time history of daisy chain resistance for pristine boards tested at 125C

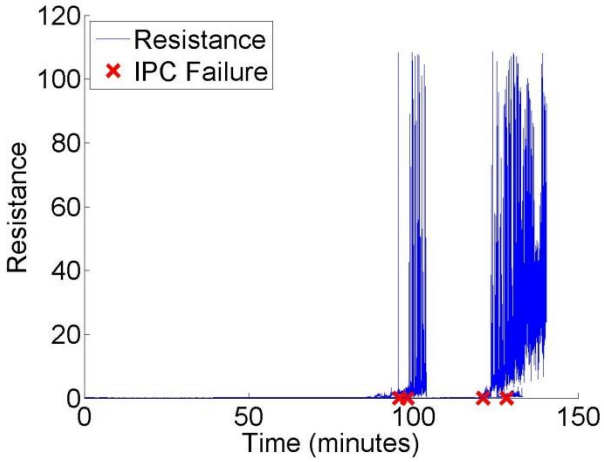


Figure 199: Time history of daisy chain resistance for isothermally aged boards tested at 25C

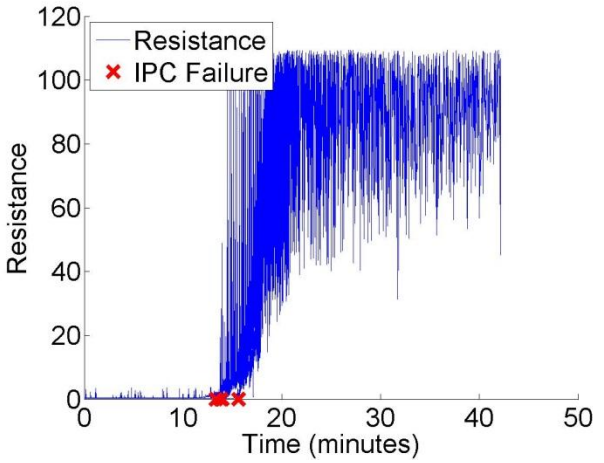


Figure 200: Time history of daisy chain resistance for isothermally aged boards tested at 125C

Table 21: Summary of fitting IPC failure data to a Weibull model

Board Condition	Test Condition	Shape Parameter	Scale Parameter	R^2
Pristine	25	2.16	68.31	0.97
Pristine	125	2.85	53.48	0.99
Aged	25	8.52	114.47	0.99
Aged	125	14.54	14.4	0.99

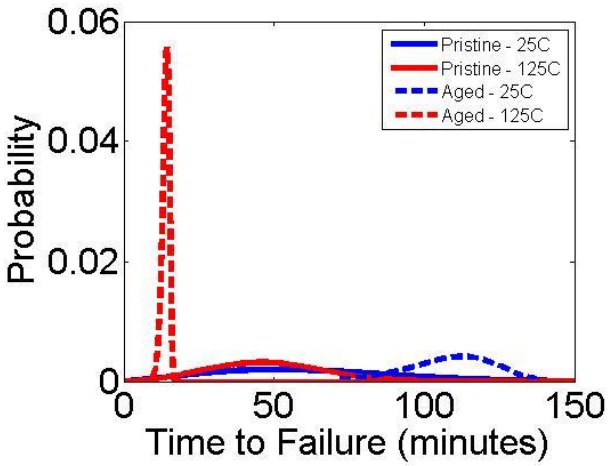


Figure 201: Graphical representation of failure distributions from extracted Weibull parameters

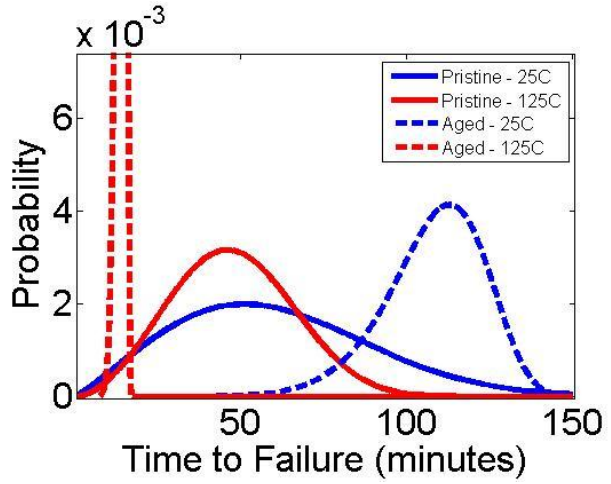


Figure 202: Zoomed view of pristine-25C, pristine-125C, and aged-25C failure distributions

The Weibull distributions highlight that the failure distributions for the two levels of pristine samples are not statistically different from each other. There is a significant difference in failure times between the aged and pristine samples. There is also a significant difference between the 25C and 125C test conditions for isothermally aged boards. If isothermal ageing in general reduces the mechanical properties of SAC solders subjected to 1-D tensions tests [Ma 2009], then the observation that aged-125C components failed before pristine components matches with the current understanding of reliability based on underlying mechanical properties. The increased time to failure for aged-25C boards is not well explained in this context. It is unclear if this result is a product of the experiment, or a result of compounding reliability issues. FR4-06 is not a certified lead free board and may be more susceptible to changes in material properties during isothermal ageing. Furthermore the effect of dynamic strains on the aged board and assembly has not been fully quantified.

12.6 Prognostic Capability

Closer inspection of the resistance time histories shows that in many cases there is a noticeable increase in resistance before the traditional definition of failure. This advanced warning of failure will be used as a leading indicator to facilitate the prediction of failure in fielded components. The resistance at which the component fails (per IPC criteria) provides a correlation between a quantity that is observable non-destructively, and the unobservable system state of the solder joint. This resistance value is termed the critical resistance for failure. Techniques borrowed from machine learning [Alpaydin 2004], such as leave one out cross

validation, will be used to help ensure that the logical fallacy known as future information leak is avoided. An example of future information leak would be using the critical resistance at which a component fails as the definition of failure for the same component. Obviously the critical resistance at failure is not observable until after failure, which makes its use a logical fallacy. It is however appropriate to use the critical resistance at failure from other tests as a guide to establish a critical resistance for the current test, and is the method implemented in this paper.

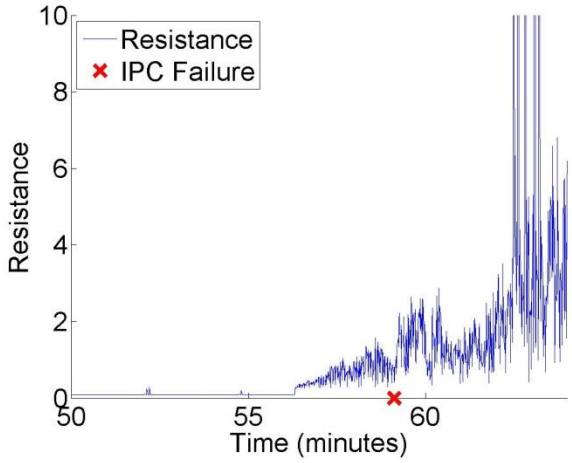


Figure 203: Zoomed view of Figure 8 at the time just before failure occurs. The resistance spectroscopy technique is sensitive to a measurable change in the resistance of the packaged before the traditional IPC definition of failure

12.7 Critical Resistance

The resistance value that coincides with failure of the solder joint is used as a failure threshold when trying to prognosticate remaining useful life. A summary of the critical resistances at failure for each of the previously tested boards is listed in Table 22.

Table 22: Summary of critical resistances for each test

		Test Number						
Board Condition	Test Condition	1	2	3	4	5	Mean Critical Resistance	Std. Dev.
Pristine	25	0.20	1.10	0.60	1.55	0.75	0.84	0.51
Pristine	125	1.05	1.32	2.00	1.60	2.10	1.61	0.44
Aged	25	0.90	0.80	0.45	1.00	2.00	1.03	0.58
Aged	125	1.00	1.00	1.52	1.95	1.50	1.39	0.40

In the prognostic algorithm presented in this paper, a single failure threshold is required. A failure threshold of 1 ohm was chosen as a conservative value.

12.8 Conclusion

The effect of simultaneous temperature and vibration testing on the reliability of BGA components have been investigated in a two level full factorial experiment. Tests at room temperature and elevated temperatures have been performed on both pristine and thermally aged components. The test conditions were designed to represent actual usage conditions that electronics deployed in harsh environments may be subjected to. Isothermal ageing had a

statistically significant impact on the reliability of tested assemblies as well as the test temperature in some of the test conditions. In addition to monitoring for failure using traditional IPC failure criteria, resistance spectroscopy measurements were made. The ability to monitor in a real time, and predict impending failures prior to the true failure were simulated. Together the monitoring, tracking and prediction capability discussed creates a complete PHM algorithm.

13 Particle Swarm Optimization for PHM

To avoid the logical fallacy of future information leak, another set of five pristine boards were run to failure in an ambient temperature of 25C. Only resistance spectroscopy measurements were recorded. Based on previous data a failure threshold of 1 ohm was used to define failure. The test will be repeatedly divided into a group of four tests, and a separate group with the test that was left out. Optimum parameters will be derived from the group of four, and then implemented on the left out data set to validate performance. This will be repeated so that each data set is left out resulting in five validation data sets. This technique is known as leave one out cross validation and is commonly used to validate machine learning algorithms. Particle swarm optimization (PSO) was the optimizer of choice due to its robust ability to cover an unfamiliar optimization space. The results presented in this section will provide insight into the performance that could be expected from a PHM algorithm on a fielded system.

13.1 Introduction

The same PBGA324 package architecture used in the simultaneous stress section was again used for this study. Five new pristine boards were mounted face down on a shaker table and run to failure during a 11gn random vibration testing sequence. During testing the full daisy chain of each board was monitored using the resistance spectroscopy measurement.

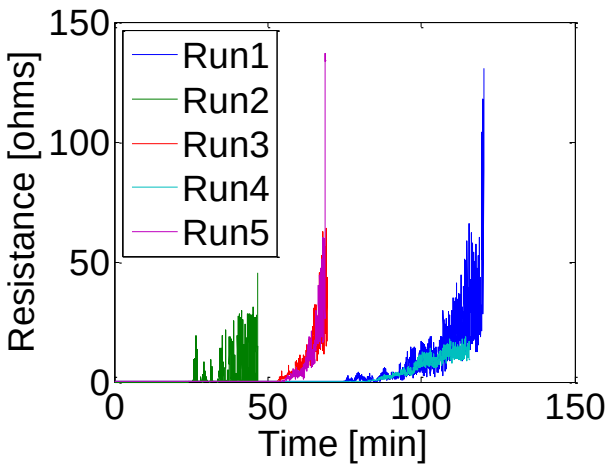


Figure 204: Additional tests run to create a test set for the PHM algorithm

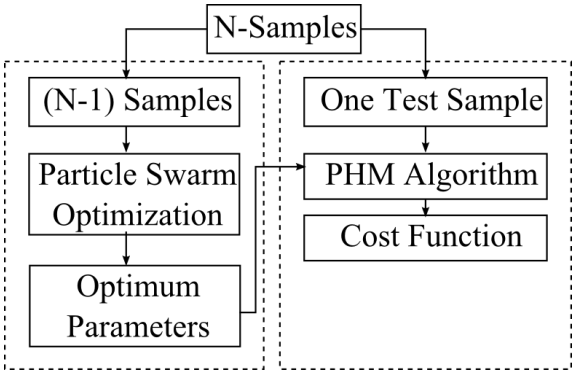


Figure 205: Leave one out cross validation procedure repeated for each data set

An extended Kalman filter (EKF) based PHM algorithm is used to facilitate the tracking of the resistance signal in a recursive manner. At a time prior to the failure of the component, the challenge is to predict when the resistance will increase to its critical value which coincides with component failure. The details of the prediction algorithm are described in detail in [Lall 2011a], but it is important to note that the algorithm requires initialization of four free parameters to function correctly. Free parameters include a process (Φ_s) and measurement noise (R) term which effect the accuracy of the EKF, and a two model terms that effect both filtering and prognostics. It is assumed that the resistance increase before failure follows an exponential curve of the form

$$R = a * exp(bt) \tag{287}$$

Where R is resistance, a is a constant model parameter, and b is a time varying parameter that is estimated inside the Kalman filter. With these free parameters many different usage conditions can be fit to the same model, but correct performance also requires successful initialization of these parameters at reasonable values. The parameters can be selected by posing the problem as an optimization. Particle swarm optimization will be used to find an optimum set of parameters to apply on the 'left out' data set used for validation. Note that the optimization of the free parameters is not influenced by the 'left out' data set.

13.2 Particle Swarm Optimization

The particle swarm optimization technique is a population based algorithm rooted in social information sharing which was originally developed to study the flight of a flock of birds, school of fish [Kennedy 1995]. PSO models social information sharing where the collaborative intelligence of the group is used to benefit each individual. Each candidate solution in PSO is called a particle. The PSO algorithm is initialized with a population of particles or candidate solutions represented by random particle locations and random particle velocities in n-dimensional solution space. The problem-solution is searched through an iterative process in the search space by adjustments in the particle locations and velocities based on prior knowledge. The particle swarm has memory capabilities while moving through the search space, which enable the swarm to remember its global best position (gbest) attained by any of the particles in the swarm and the corresponding solution or fitness. In addition each particle remembers their personal best position (pbest) and the associated fitness value attained while moving through the solution space. PSO has been shown to solve non-linear, differentiable, multimodal optimization problems [Krohling 2004, Clerc 2002]. Each particle evaluates its fitness value and compares the

fitness value with its best position. If the current fitness value is better than pbest, the particle sets the current fitness value to pbest and saves the current location in the n-dimensional space to the pbest location. PSO also compares the particle's fitness evaluation to the swarm's overall previous best, gbest. If the particle's current fitness value is better than gbest, then the gbest is reset to the current particle array index and fitness value. In this paper the fitness values is evaluated by the cost function:

$$J(a, b, \Phi_s, R) = 1 - \frac{1}{N}\sum_{k}\left(C_\beta * \beta_k + C_{RA} * RA_k\right) \tag{288}$$

Where, β is a PHM metric that quantifies the precesion of predictions, and RA is the relative accuracy of each prediction. PHM metrics are represented as vectors containing a value between zero and one for each prediction time step and N is the total number of predictions. C_β and C_{RA} are weights that sum to a value of unity. In this case $C_\beta = C_{RA} = 0.5$, which indicates an equal importance of precision and accuracy. The first two inputs, a and b, are parameters used in the state propagation model, equation (287). The other two respective inputs, Φ_s and R, are the process noise and measurement noise terms from the Kalman filter. To calculate the cost function the specified free parameters are input into the PHM algorithm and run on a data set. Predictions are simulated as if they were being made in real time. Standard performance metrics are used to quantify the success of the predictions. The result of the cost function is a number between 0 and 1, where a lower score is better. PSO will aim to find the optimum parameters to minimize the cost function, and as a result maximize performance of the PHM algorithm. A particle swarm optimization starts by seeding particles at random in the feature space. Two of the four free parameters being optimized are shown in Figure 206. Initial particles are chosen to span many orders of magnitude, so all calculations are performed in the log space to avoid biasing the

swarm metrics with very large or very small numbers. The position of the i[th] particle in n-dimensional space is represented by:

$$x_i = \lfloor x_{i1} \quad x_{i2} \quad \cdots \quad x_{in} \rfloor^T = \lfloor a_i \quad b_i \quad \Phi_{s,i} \quad R_i \rfloor^T \tag{289}$$

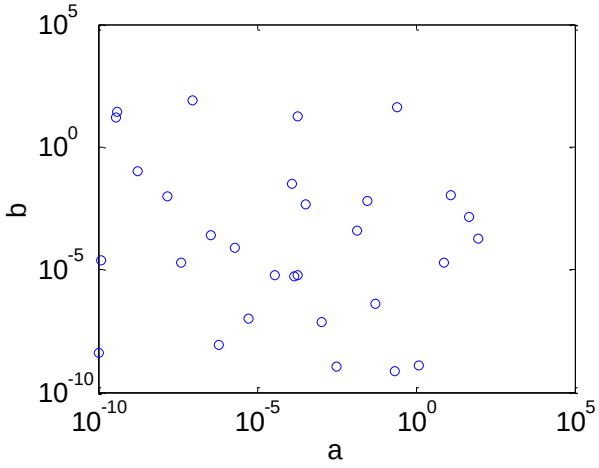


Figure 206: Initialization of particles for two of the free parameters

The velocity of the i[th] particle in n-dimensional space is represented by:

$$v_i = \lfloor v_{i1} \quad v_{i2} \quad \cdots \quad v_{in} \rfloor^T \tag{290}$$

Location of the particle's previous best position based on the obtained fitness value in n-dimensional space is represented by:

$$p_b = \lfloor p_{b1} \quad p_{b2} \quad \cdots \quad p_{bn} \rfloor^T = \lfloor a_{bi} \quad b_{bi} \quad \Phi_{s,bi} \quad R_{bi} \rfloor^T \tag{291}$$

The index g represents index of the best particle among all particles in the swarm. This term is also known as the global best position of the swarm:

$$p_g = \lfloor p_{g1} \quad p_{g2} \quad \cdots \quad p_{gn} \rfloor^T = \left\lfloor a_g \quad b_g \quad \Phi_{s,g} \quad R_g \right\rfloor^T \tag{292}$$

In the original Kennedy [1995] formulation, velocity of the particle is changed according to the equation:

$$v_i(t+1) = v_i(t) + w_1\phi_1[p_b - x_i(t)] + w_2\phi_2\left[p_g - x_i(t)\right] \quad (293)$$

Where $v_i(t)$ is the velocity of the i[th] particle at time t, $v_i(t+1)$ is the velocity of the i[th] particle at time t+1, ϕ_1 and ϕ_2 are random numbers between 0 and 1, w_1 and w_2 are the scaled weights for the cognitive and social terms respectively. Kennedy [1995] specifies a value of 2 for both the scaled weights w_1 and w_2. The second term in equation (293) is the cognitive term which represents the independent behavior of the particle itself. The third term in equation (293) is the social term which represents the collaboration between the particles. The weightsw_1 and w_2 are used to adjust the weighting of the social and the cognitive parts that pull the particles towards the pbest and gbest positions. Krohling's [2004] formulation has been used in this paper, which involves using a Gaussian distribution of random numbers (ϕ_1 and ϕ_2) instead of a uniform distribution of random numbers which is used in standard PSO. The larger standard deviation provided by the Gaussian distribution compared to the uniform one, improves the ability of the PSO algorithm to escape from local minima.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (294)$$

Where f is the probability density function representing a Gaussian distribution for random variables x, σ is the standard deviation, and μ is the mean value of the random variable distribution, which has been set to zero for this optimization. The following form of the equation (28) has been used for PSO:

$$v_i(t+1) = v_i(t) + |\phi_1|[p_b - x_i(t)] + |\phi_2|\left[p_g - x_i(t)\right] \quad (295)$$

Where ϕ_1 and ϕ_2 are positive random numbers generated according to the absolute value of the Gaussian distribution. The static scaled weights have been omitted in this formulation. The position of the particle is then changed as follows:

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{296}$$

Since this present optimization is 4-dimensional, and surfaces of greater dimensionality do not allow for easy visualization, a set of metrics has been used to assess the status of the swarm during the simulation. The swarm metrics have been observed over time as the swarm is allowed to iterate through the search process. The swarm metrics used include:

(1) global best position (gbest)
(2) swarm standard deviation or the spread of the swarm during any iteration
(3) distance moved by the global best position.

This has been calculated from the particle positions as follows:

$$\sigma_{swarm} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - \bar{x})^2} \tag{297}$$

Where m is the number of particles in the swarm, x_i is the position of the i[th] particle, $\bar{x}$ is the mean of the particle positions in the n-dimensional space for m-particles and is computed as follows:

$$\begin{aligned} x_1 &= \lfloor x_{11} \quad x_{12} \quad \cdots \quad x_{1n} \rfloor^T \\ x_2 &= \lfloor x_{21} \quad x_{22} \quad \cdots \quad x_{2n} \rfloor^T \\ &\vdots \\ x_m &= \lfloor x_{m1} \quad x_{m2} \quad \cdots \quad x_{mn} \rfloor^T \end{aligned} \tag{298}$$

$$\bar{x} = \frac{1}{m}\left\lfloor \sum_{j=1}^{m} x_{j1} \quad \sum_{j=1}^{m} x_{j2} \quad \cdots \quad \sum_{j=1}^{m} x_{jn} \right\rfloor^T \tag{299}$$

In the present case, n = 4, corresponding to the variables: a, b,Φ_s, and R. The distance travelled by the global best position is calculated as follows:

$$p_g(t) = \lfloor p_{g1}(t) \quad p_{g2}(t) \quad \cdots \quad p_{gn}(t) \rfloor^T \tag{300}$$

$$\Delta p_g(t) = \sqrt{\left(p_{g1}(t+1) - p_{g1}(t)\right)^2 + \left(p_{g2}(t+1) - p_{g2}(t)\right)^2 + \cdots + \left(p_{gn}(t+1) - p_{gn}(t)\right)^2} \tag{301}$$

$$\Delta p_g(t) = \sqrt{\sum_{j=1}^{n} \left(p_{gn}(t+1) - p_{gn}(t)\right)^2}$$

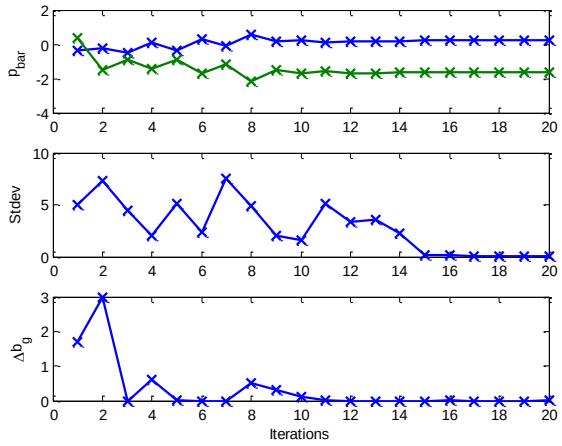


Figure 207: Time history of swarm metrics. PSO terminates when the metrics converge below a specified tolerance.

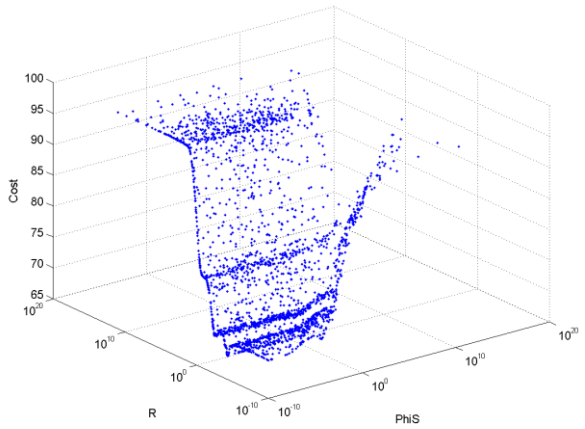


Figure 208: Search space of the PSO algorithm visualized for the process noise and measurement noise terms in the Kalman filter

For the parameters a and b, the time history of the swarm metrics is shown in Figure 207.

Each of the four data sets in the training set are optimized individually using PSO. Then the

optimized parameters are averaged to obtain a single set of parameters to pass to the PHM algorithm used in the test/validation stage. The search space for a typical data set is shown in Figure 208. From the plot it can be seen that it is important to choose the measurement noise term correctly. A comparison of optimizing the free parameters by hand, and using the PSO algorithm is shown in Figure 209.

Table 23: Results of leave one out cross validation

Validation Run	a	b	Φ_s	R	cost
1	1.45	12.42	0.10	0.06	0.66
2	0.76	12.55	0.12	0.06	0.78
3	1.59	9.64	0.07	6.73E-04	0.50
4	1.67	14.31	0.12	0.06	0.77
5	1.34	9.64	0.06	0.06	0.96
			Average Cost: 0.73		

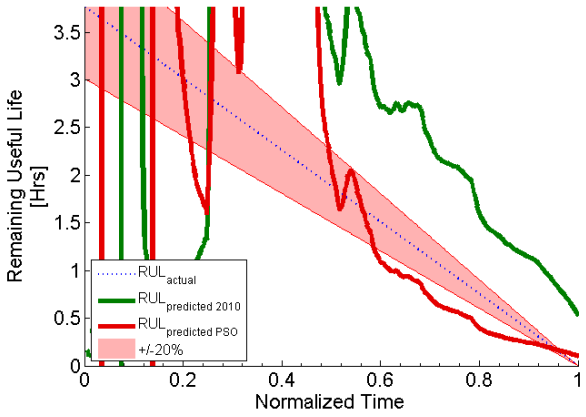


Figure 209: Comparison of PSO method against previously published results to show the improvement using the new method.

Unfortunately optimizing a PHM algorithm using the full data set is a future information leak. Therefore to demonstrate the performance one could expect from the algorithm on a data

set not used to train the algorithm the data set was divided two parts. The first N% of the data set was used as an input to the PSO algorithm to determine optimum parameters. Then the performance of the algorithm using optimum parameters derived from the initial N% of the data set on the remaining (100-N)% of data was used as a test. Figure 210 shows that only a small portion of the full data set is needed to derive optimum parameters. After the signal reaches approximately 25% of the failure threshold the PSO algorithm converges to an optimum value for the free parameters in the Kalman filter and results in the best possible performance of the PHM algorithm. The choice of 1 ohm as the failure threshold for this data set was based on previous data described in the temperature/vibration section. Cross validation prediction results shown in Figure 210 provides insight into the performance that could be expected if the PHM algorithm was implemented in a functional product.

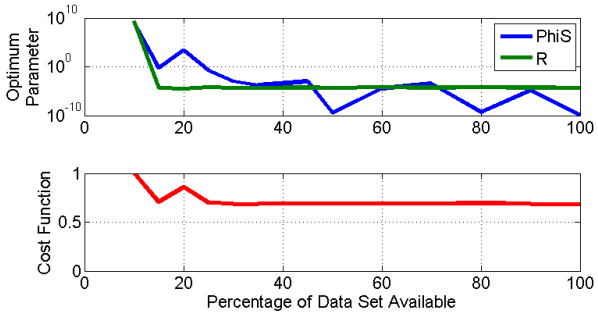


Figure 210: Performance of PHM algorithm when using only portion of the initial data set to find optimum parameters

13.3 Conclusion

Leave out cross validation was utilized to evaluate the expected performance of the PHM algorithm as if it were implemented in a functional system. Results from the run-to- failure tests shown in this section highlight the robustness and repeatability of the described techniques. Particle warm optimization was used to find an optimum set of free parameters that would

quickly converge to an acceptable configuration for prognosticating failure in safety and mission critical electronics.

14 Cost justification for Implementing PHM on circuit boards

Prognostic health management is a method for assuring the reliability of a system by monitoring the system in real time as it is used in the field. As the system wears out, but before failure, information that facilitates decision making about the future use of the system is delivered to the user. In this section, a cost justification has been developed for incorporating the additional circuitry needed to enable prognostics for electrical components directly onto a functional circuit board. Implementing PHM capability for circuit boards will add additional cost to a system, so high reliability systems where the cost of failure is high are easiest to cost justify for prognostics. Aerospace, defense and automotive, applications qualify as high reliability systems. Application domains that require high uptime, minimal amounts of unplanned maintenance, and controllable operating costs can also be cost justified for prognostics since they can benefit from the pro-active management of failures facilitated by PHM. Intangible criteria such as safety or the cost of human life also motivated the need for PHM, so often time's projects are labeled strategic, and not subjected to the discipline of a financial analysis. This paper will show rigorous methods for assessing the decision to invest in PHM for electronics. The uncertain nature of research and development (R&D) and difficult to predict future economic conditions is not well captured by traditional discounted cash flow (DCF) methods. An approach known as the Datar-Mathews (DM) method will extend the DCF methods to be equivalent to a real options analysis and the Black-Scholes formula. The DM method is intuitive and uses concepts familiar to most engineers and technical managers.

14.1 Introduction

Prognostic health management for electronics has reached a technology readiness level where proof of concept demonstrations exist [Lall 2010[a], 2011[a-c]]. Previously technical information about the sensors [Lall 2009a-c], prediction algorithms [Orchard 2007, Goebel 2008, Saha 2009, Lall 2010[a], 2011[a-c]] and risk based decision making techniques [Engle 2009, Lall 2010[a]] that combine to form a complete PHM framework for BGA components were published. A long history of the study of failure mechanics and leading indicators of failure in electronics under a variety of loads and conditions [Lall2004[a-d], 2005[a-b], 2006[a-f], 2007[a-e], 2008[a-f]] has provided the foundation for the most recent work that closely embodies the PHM methodology. To date all proof of concepts have used jumper wires to a connect circuit boards under test to custom circuitry sitting on a lab table to enable prognosticating system health (predicting future failures based on measured leading indicators of failure, in real time). While this is sufficient for a proof of concept, it is not in the form of a functional product. A complete PHM implementation would embed all necessary circuitry onto the board being monitored while maintaining the same form factor. From a system level there would be no difference in the PHM enabled board and the non-PHM enabled version, except for the availability of prognostic health information on output pins. Currently there is no method for modifying an existing circuit board to facilitate real time, non-destructive, computationally and financially tractable health monitoring techniques. This paper proposes a method for cost justifying the research and development costs needed to convert an existing proof of concepts into an embedded PHM solution on a new circuit board. To handle the financial uncertainty of an investment that depends on hard to predict future economic conditions a modified Black-Scholes formula is used. The presented approach [Mathews 2007] is an algebraically equivalent formulation to the traditional Black-Scholes formula. Motivation for

using the DM method in place of the Black-Scholes formulation is discussed. Many of the concepts used in the DM method will be familiar for engineers and technical managers.

It is important to distinguish the approach presented in this paper as a bottom-up PHM cost justification, as opposed to the top-down approach that has been extensively studied [Byer 2001, Wilkinson 2004, Goodman 2005, Carraco 2006, Banks 2007, 2009, Sandborn 2007, Feldman 2009, Luna 2009]. A top-down approach starts with high level requirements like safety, availability and operating cost of a large system (airplane, car, turbine, etc.) and argues that savings can be realized if prognostic technology that predict failures were to be implemented. Often times it is brazenly assumed that the system can be characterized to the granularity where PHM can be efficiently implemented on components. Many of these approaches do not delve into technical details, or linger on the possibility that the technology for the level of insight they assume does not exist. Top-down analysis tend to be overly general, or overly application specific. In contrast, bottom-up approaches start with a single component (in this paper, part of a circuit board). A cost justification is made for the component to determine if it is beneficial to invest in the development of PHM for that specific component. Both approaches are appropriate in certain situations, but the bottom-up PHM cost justification is less studied because it requires a prototype as a starting point.

To motivate this work, Figure 211 shows the price index for agricultural inputs (land, fuel, feed, etc.) compared against the cost index of agricultural outputs (grain, dairy, wheat, etc.). The relative price of agricultural inputs has increased in pace with the overall US price index. The cost of food has remained relatively steady. This gap can be attributed to improved efficiencies in the agriculture industry. From this data it can be seen that now more than ever there is a significant opportunity cost associated with unplanned downtime and lost productivity.

Modern agricultural equipment, like most modern systems, have multiple critical subsystems that rely on electronics to perform correctly.

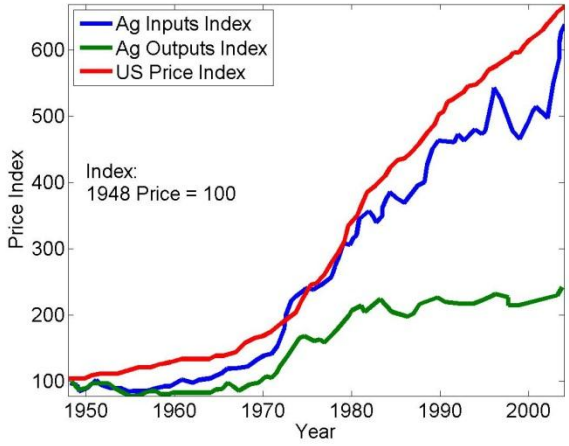


Figure 211: Price index for 1948-2004 showing trends in agriculture inputs and outputs. Re-drawn from Fuglie, Keith O., James M. MacDonald, and Eldon Ball. Productivity Growth in U.S. Agriculture. EB-9, U.S. Dept. of Agriculture, Econ. Res. Serv. September 2007.

Many other sectors such as aerospace, automotive and the military also depend heavily on electronic subsystems performing an expected function at a certain time. Electronic rich systems are becoming increasingly common and this work attempts to quantify if it is beneficial to implement PHM for electronics into a particular system. Incorporating PHM capability into a new circuit board design is relatively straight forward. Incorporating PHM capability into an existing board design requires a reversion of the board, which can be an expensive hurdle to overcome.

14.2 The Real Option for R&D Investment in Prognostic Technology

A real option is a class of problems where an investor has the option, but not obligation, to purchase an investment at a set value at some time in the future. Of course you must pay for such an option, so the challenge is to determine if given the future uncertainty of the investments value, is it beneficial to pay for an option today that may or may not be valuable to exercise at a

later date. This problem was famously solved for financial options [Black 1973]. Real option analysis uses the analogy of financial options which are traded on exchanges as commodities as a framework for making business decisions and capital investments. Real option analysis has been used in engineering economics problems relating to maintenance decision making [Heredia-Zavoni 2004]. A typical net present value analysis is not well suited for this class of problems and will undervalue the true value of the real option to perform maintenance.

In the terminology used in the real option literature, investing in R&D for prognostics gives you the right, but not obligation, to purchase PHM technology at below market value in the future. The purchase price is represented as the launch price for the product. It goes without saying that you would not launch the product if you did not think you would make money. Real options analysis helps quantify the probability that you will make a certain payoff if you launch the product, so you can decide if the R&D expenditure is a prudent investment in advance.

14.3 Black-Scholes Formula

A standard real option analysis uses the Black-Scholes formula (BSF). The approach presented in this paper [Mathews 2007] is algebraically equal to the BSF, but has the benefit of representing uncertainty with easier to define values. A Black-Scholes analysis assumes that the value of an investment in the future is a function of time, t, and volatility, σ. The closed form solution for valuing a financial option is the BSF [Hull 2000],

$$C_{BSF} = S_0 N(d_1) - K e^{-R_f \Delta T} N(d_2) \tag{302}$$
$$d_1 = \frac{\ln(S_0 / K) + (r + \sigma^2 / 2)\Delta T}{\sigma \sqrt{\Delta T}}$$
$$d_2 = d_1 - \sigma \sqrt{\Delta T}$$

Where C_{BSF} is the value of an option known as a European call option. S_o is the current price of the asset, N(x) is the cumulative probability density function for a standard normal distribution, K is the strike price or exercise price, R_f is the risk free rate of return compounded continuously. The BSF implicitly assumes that the future price of the asset can be modeled as a stochastic process (aka Brownian motion, Figure 212) and that the volatility of the underlying process is constant. Volatility can be easily calculated for many financial options using historical trading records that are readily available. When valuing a real option, as opposed to a commodity, the analogy of volatility may not have much physical meaning. Brownian motion of an investment can be modeled in recursive form as [Higham 2004]

$$S_t = S_{t-1} \exp\left(\left(\mu - 0.5\sigma^2\right)\Delta t + \xi\sigma\sqrt{\Delta t}\right) \tag{303}$$

Where S_t is price at time step t, μ is the growth rate, Δt is the length of the time step, and ξ is a normal random number drawn from N(0,1).

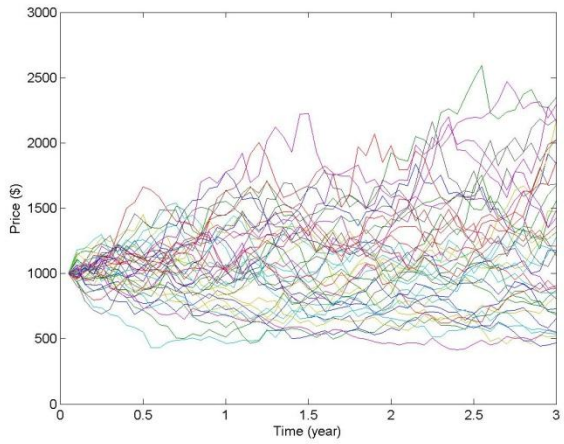


Figure 212: Hypothetical evolution in time of an investment using Equation (1). In this example the volatility is 50%. Discretely simulated values have been connected with lines to show trends

For this relatively simple case a closed form solution of the distribution of possible future prices, denoted as S, that an investment would be expected to be worth at the end of the

investment time window is available. S_o represents the starting price and ΔT is the length of the total investment window.

$$f(S) = \frac{\exp\left[\frac{-\left(\ln\left(\frac{S}{S_o}\right)-\left(\mu-0.5\sigma^2\right)\Delta T\right)^2}{2\sigma^2\Delta T}\right]}{S\sigma\sqrt{2\pi\Delta T}} \tag{304}$$

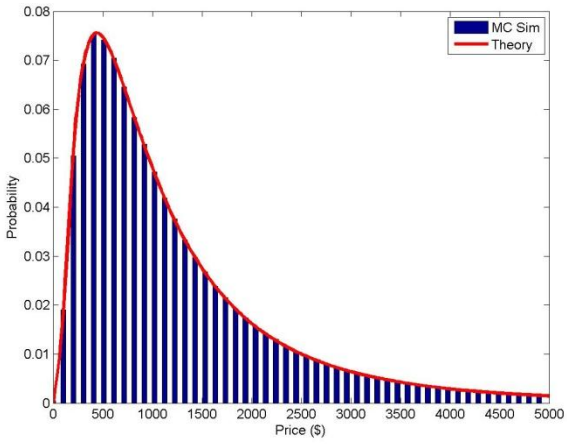


Figure 213: Comparison of a Monte Carlo simulation of the future price of an investment compared to the analytical result. The analytical solution is a lognormal distribution. The Monte Carlo simulation used 106 trials of Equation (1) binned at 60 locations.

14.4 Datar-Mathews Approach

The Datar-Mathews (DM) approach is an alternative method for valuing a real option. Under the right set of assumptions, it can be shown to be algebraically equivalent to the Black-Scholes formula. The DM approach is useful because it can be used in situations where application of the BSF may be impractical.

The DM method can be represented as

$$C_{DM} = E\left[\max\left\{e^{-R_r t}\bar{S} - e^{-R_f t}K\right\},0\right] \tag{305}$$

Where C_{DM} is the value of the European call option, E represents and expectation or mean. The max operator takes the maximum between the value in the curly brackets and zero. Two different interest rates are utilized in this formula. R_f is the previously defined risk free rate of return. R_r is the risky asset discount rate, commonly called a hurdle rate. Cash flows subjected to market risk can be discounted by the risky asset rate. Cash flows controlled by less risky sources, such as management, can be discounted with the risk free rate of return. The upcoming example will expand on the differences and applications of R_f and R_r. The $\bar{S}$ term represents a distribution of possible future values of an asset. There are no limitations on how the $\bar{S}$ term is obtained. It can be a closed form distribution or obtained through any other simulation method such as Monte Carlo.

14.5 Equivalence of Black-Scholes and Datar-Mathews Approach

Under a set of assumptions, the Black-Scholes and Datar-Mathews approaches can be shown to be algebraically equivalent [Datar 2004]. Take for example the business decision to invest in research and development (R&D) to commercialize a new technology. This example is a simplified example of the decision to invest in R&D to develop PHM capability for electronics.

If you believed that the commercialization of a technology could realize an operating profit with an investment in R&D in three years you could frame the option to invest in R&D as a real option. In year zero you would invest in the research. In year two if the R&D was successful, you would have the option to launch the product. There would be a launch cost to buy tooling and start production of your new product. You could also expect to earn an operating profit in year two. There are no certainties in R&D, so you judge that you have a 50% chance of

success. The value of the option to launch your product at year two determines how much money you should spend on R&D in year zero.

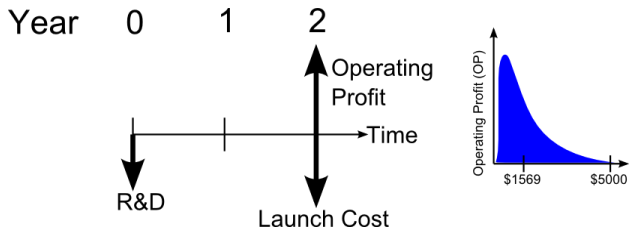


Figure 214: Cash flow diagram for example problem

You model the expected future operating profits as a lognormal distribution with a mean (OP_{mean}) of \$1569 and a standard deviation of \$1651. The launch cost is estimated as \$500. For the purposes of using the BSF, your volatility is 50%. The risk free rate of return (R_f) is taken as the cooperate bond rate, 5%. The risky asset discount rate (R_r) is taken as your corporations weighted average cost of capital, 15%. The time frame for the investment is three years.

Asset (S)	$e^{-R_r t}(OP_{mean})$	1000
Exercise (K)		500
Volatility (σ)		50%
d1, d2		1.4069, 0.5409
N(d1), N(d2)		0.9203, 0.7057
Option Value (C_{BSF})		616.84

This result implies that you should not expend more than \$616 on research and development. Similarly the DM method can be used to value the option of investing in R&D.

Asset	Make N draws from the OP distribution	$\bar{S}$
Exercise (K)		500
Pay Off	$\max\left\{\left(e^{-R_r t}\bar{S} - e^{-R_f t}K\right), 0\right\}$	
Option Value (C_{DM})	E[pay off]	616.84

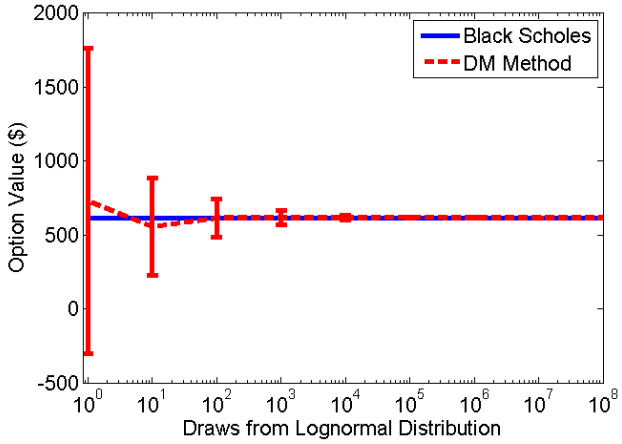


Figure 215: Convergence of the DM method to the Black-Scholes value for the example problem. Error bars represent one standard deviation determined from 30 different trials.

Figure 215 shows that the DM method converges quickly to the closed form solution obtained using the BSF. When assuming a lognormal distribution of future prices for an asset the BSF and DM methods give the same result. The advantage of using the DM method is that the approach is generic enough to use models that describe the future value of an asset in alternative methods. In practice, having knowledge of the volatility of an investment, especially a capital investment such as the payout on a circuit board with PHM capability is very difficult to quantify. PHM capability for electronics is not a traded commodity, and there is currently no historical data to determine volatility since no commercialized products exist. It would be

reasonable to match the volatility of an investment in an unknown situation with the known volatility from a company operating in a different sector where product development is risky, such as a bio-technology firm, but it would be difficult to validate your choice since it not based on a model. An alternative method to determine the value of the option to invest in PHM technology is to use the Monte Carlo pricing approach and relax the assumption that the value of the investment follows a Brownian motion model and that volatility is constant over the investment time frame. An alternative model could use operating profit projections to predict future values of the investment instead of the Brownian motion model. Operating profit projections are intuitive approach to thinking about the future value of an investment and standardized methods for creating predictions in your specific market sector probably already exist. This alternative approach will be used in this paper to assess the value of implementing PHM capability into a circuit board.

Because it is widely introduced in introductory finance classes, the net present value will be used to will be used to discount cash flows in the remainder of the paper. The net present value (NPV) is defined as [Higham 2004].

$$NPV = \sum_{n=0}^{\Delta T} \frac{CF_n}{(1+r)^n} \tag{306}$$

Where n is the year, ΔT is the total number of years in the investment window, CF_n is the cash flow for year n, and r is the annual discount rate. Note the subtle shift from continuously compound interest rates used in the BSF to annually compounded interest because of their familiarity. As long as units are consistent, the DM method can use either type of interest rate. Annually compounded interest rates will be denoted with lower case variable names to provide a distinction. In an NPV calculation the discount rate is usually taken as the risk free rate of return.

In real options analysis some cash flows are not dependent on market risk, and can be assigned a different discount rate. For example the cash required to launch a new project is controlled by management, and not the market, and therefore is not as risky.

Using the language of discounted cash flows helps ease the transition to real options thinking because of the similarity to more commonly used financial analysis. Furthermore the method can help bridge the gap between engineers, engineering managers and financial managers. Engineers are involved in the R&D and manufacturing, but probably have limited training in financial analysis. Financial managers are less familiar with the technical details of the engineering but have the appropriate financial training. To make the best possible business decision both groups need to collaborate effectively, and the DM method can provide a bridge to successful collaboration.

14.6 Calculating the Value of the Option to Invest in PHM

The real option is formulated as:

If I invest in R&D to develop PHM capability for a circuit board, what will the option, but not obligation, to build PHM embedded circuit boards be worth given future uncertainties?

Research and development costs would involve building prototypes for testing, and electrical engineering design to layout and route test circuitry that will be incorporated into the boards to facilitate PHM. Accelerated life testing will be used to quickly stress prototypes to failure and test the success of a PHM implementation. Multiple iterations will be needed to develop the technology to a sufficiently mature level.

Description	Quantity	Unit Cost	Extended Cost

EE Design	10 wks.	$ 100 / hr.	$40,000
PCB Builds	5 builds	$2,000 / build	$10,000
Accelerated Testing	9 Months	$ 5,000 / month	$45,000
		Total R& D Costs	$95,000
Assumptions: Components and circuit boards to build 60 assemblies assume raw material costs of $1000, plus an additional $1000 in labor for layout and assembly time. Two weeks of EE Design per build.			

Consider the following scenario. In year zero an investment in R&D is made to increase maturity of existing PHM proof of concepts. In year one, if the PHM technology matures sufficiently during the R&D phase it will be incorporated into production boards which will require a re-design of the board and incur large project launch costs. If at the end of the R&D phase the outlook for sales is revised downwards, the relatively small R&D investment becomes a sunk cost that cannot be recovered. In this example, year one is the decision point for the project and we desire to know the value of the option to terminate the project. Failure to sufficiently mature the technology is an example of an event that would revise sales downward. Adding uncertainty to the analysis helps address the question: what is the risk adjusted payoff, knowing that in some cases the project may fail and the R&D investment will be lost. The launch cost, or cost of revising and re-tooling an existing PCB design is estimated at $1M dollars and will be incurred as the project launch cost in year one. By year two production will be at full speed, and profit modeling can generate a scenario of pessimistic/likely/optimistic operating profits.

Year	0	1	2	3	4	5	6
Pessimsitic ($)			(597,918)	(557,710)	(513,481)	(464,829)	(411,312)
Most Probable ($)			161,195	275,374	406,680	557,682	731,335
Optimistic ($)			921,857	1,166,229	1,459,474	1,811,369	2,233,644

Figure 216: Three predicted levels of operating profits. These numbers are derived from a more detailed analysis shown in the appendix.

Each scenario represents a potential future reality. In the pessimistic case unit costs are high and poor results from the R&D phase result in a low technology diffusion level. In the optimistic case year to year growth is better than expected as the technology is quickly adopted and costs are lower than expected. A full derivation of each scenario is shown in the appendix.

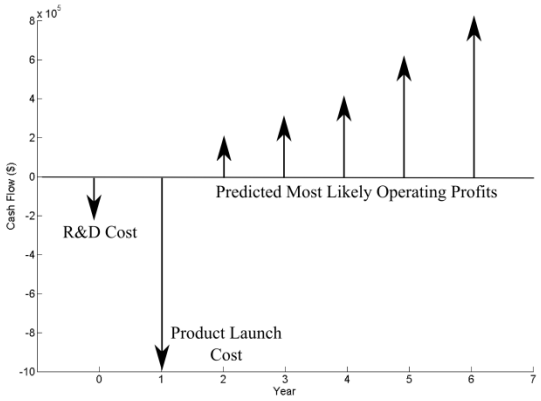


Figure 217: Cash flow diagram for the most probable set of operating profits

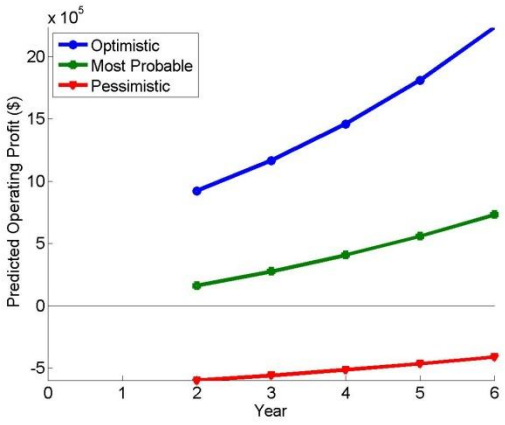


Figure 218: Three levels of predictions for the future operating profits

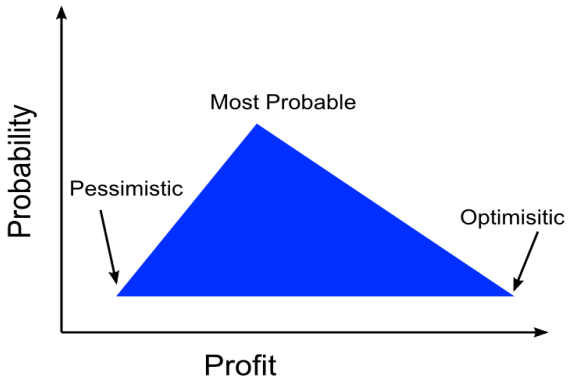


Figure 219: The triangular distribution which approximates more complicated types of distributions. Since values are uncertain, assigning values more accurate than high/medium/low is probably overly thorough

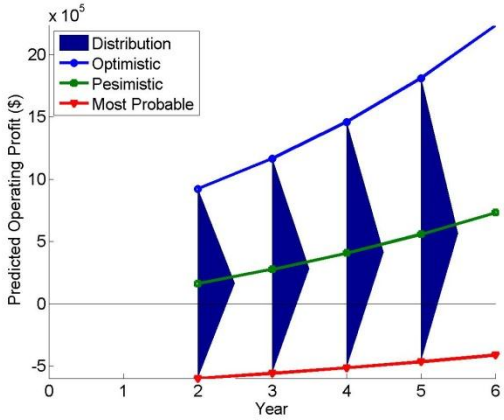


Figure 220: Triangle distributions for each year superimposed on the operating profits graph

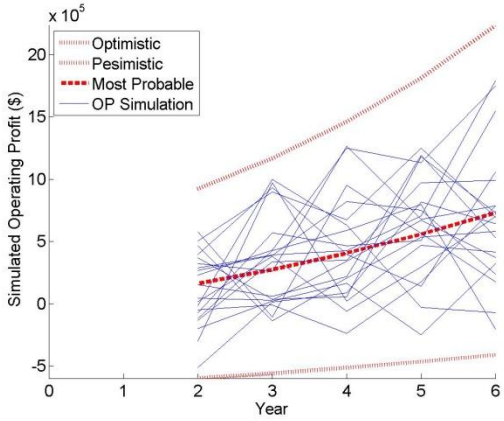


Figure 221: The first 10 (of many) operating profit simulations

Due to its intuitive benefits in practical situations a triangle distribution, which roughly approximates a lognormal distribution, is used to represent future values, but any distribution can be used with this method. A Monte Carlo simulation is used to simulate thousands of scenarios where discrete profits are drawn from the defined probability distributions at random. For each run of the simulation a discrete operating profit (OP) is generated for each year based on the pessimistic/likely/optimistic predictions. Draws between successive years were 80% correlated since it would be unlikely to move from the extreme upper end of a profit prediction to the extreme lower end of a prediction between years. Then the present value is calculated with the annual risky asset rate of return, r_r over the period of the investment horizon ΔT .

$$PV_k^{OP} = \sum_{n=0}^{\Delta T} \frac{OP_n}{(1+r_r)^n} \qquad \text{where k = 1,2,3,…N} \tag{307}$$

The present value of the launch cost, LC, is a function of its less risky discount rate, such as the corporate bond rate $r_{f,}$ over the period of time to the decision point ΔT_d.

$$PV_k^{LC} = \frac{LC}{(1+r_f)^{\Delta T_d}} \qquad \text{where k = 1,2,3,...N} \tag{308}$$

The real option for each run is evaluated as either zero in the case when at the decision point the project is unfavorable and is cancelled ,and the R&D cost becomes a sunk cost. Or in the case where at the decision point the project is favorable and the project is launched. The value of the option, RO, for each Monte Carlo run, k, is calculated as:

$$\text{RO}_\text{k} = \max(\{PV_k^{OP} - PV_k^{LC}\},0) \quad \text{where k = 1,2,3,...N} \tag{309}$$

Where PV_k^{OP} is the present value of the discrete operating profit simulation, and PV_k^{LC} is the present value of the launch cost. Performing many runs of the Monte Carlo simulation

allows the range of possible scenarios to be investigated. Figure 222 examines two hypothetical runs out of many runs in a Monte Carlo simulation to demonstrate how the value of the option to launch a product is calculated.

	Run 1	Run 2
Simulated discrete operating profit ($) (Random Draws)	Yr. 2: -500,000 Yr. 3: -300,000 Yr. 4: 100,000 Yr. 5: 300,000 Yr. 6: 500,000	Yr. 2: 100,000 Yr. 3: 500,000 Yr. 4: 800,000 Yr. 5: 1,000,000 Yr. 6: 1,250,000
Discount Rate	The risky asset rate of return is taken as 15% and used to discount operating profits. The cash for the product launch is controlled by management, not the market, and is discounted at 5%. Note that year zero counts as the first year in the investment window.	
Present Value of Launch Cost	$PV_1^{LC} = \frac{1000000}{(1+0.05)^1} = \$952381$	$PV_2^{LC} = \frac{1000000}{(1+0.05)^1} = \$952381$
Net Present Value of discrete operating profit	$PV_1^{OP} = \frac{-500000}{(1+0.15)^2} + \frac{-300000}{(1+0.15)^3} + \frac{100000}{(1+0.15)^4} + \frac{300000}{(1+0.15)^5} + \frac{500000}{(1+0.15)^6} = -\$153000$	$PV_2^{OP} = \frac{100000}{(1+0.15)^2} + \frac{500000}{(1+0.15)^3} + \frac{800000}{(1+0.15)^4} + \frac{1000000}{(1+0.15)^5} + \frac{1250000}{(1+0.15)^6} = \$1.9x10^6$
Project Net Present Value	$NPV_1 = PV_1^{OP} - PV_1^{LC} = -\$1.79x10^6$	$NPV_2 = PV_2^{OP} - PV_2^{LC} = \$947000$
Real Option Value	$RO_1 = max\{0, NPV_1\} = \$0$	$RO_2 = max\{0, NPV_2\} = \$947000$
Success/ Failure	In this scenario the operating profits are not favorable. Do not launch the product. Project is a failure and the R&D investment would be lost.	In this scenario the operating profits are favorable and the option to launch the product is valuable. Project is a success.

Figure 222: Two hypothetical runs of a Monte Carlo simulation to illustrate the simulation method

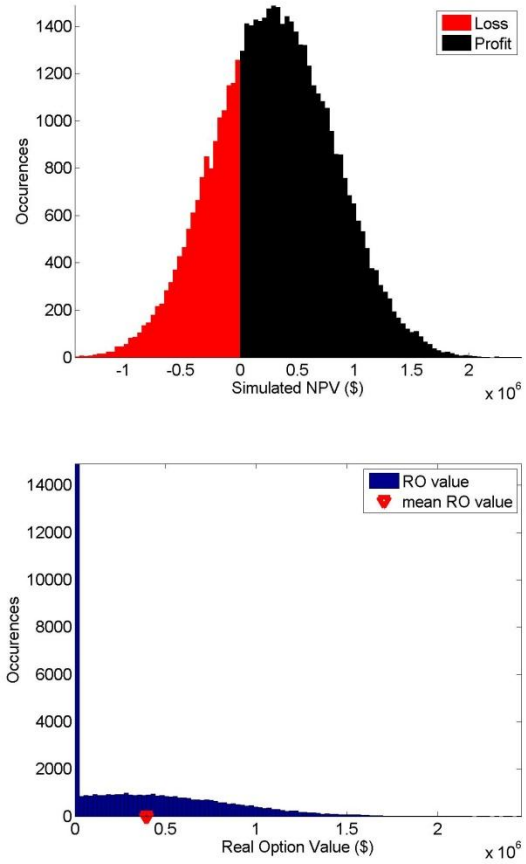


Figure 223: Results of Monte Carlo simulation used to value the option of investing in R&D

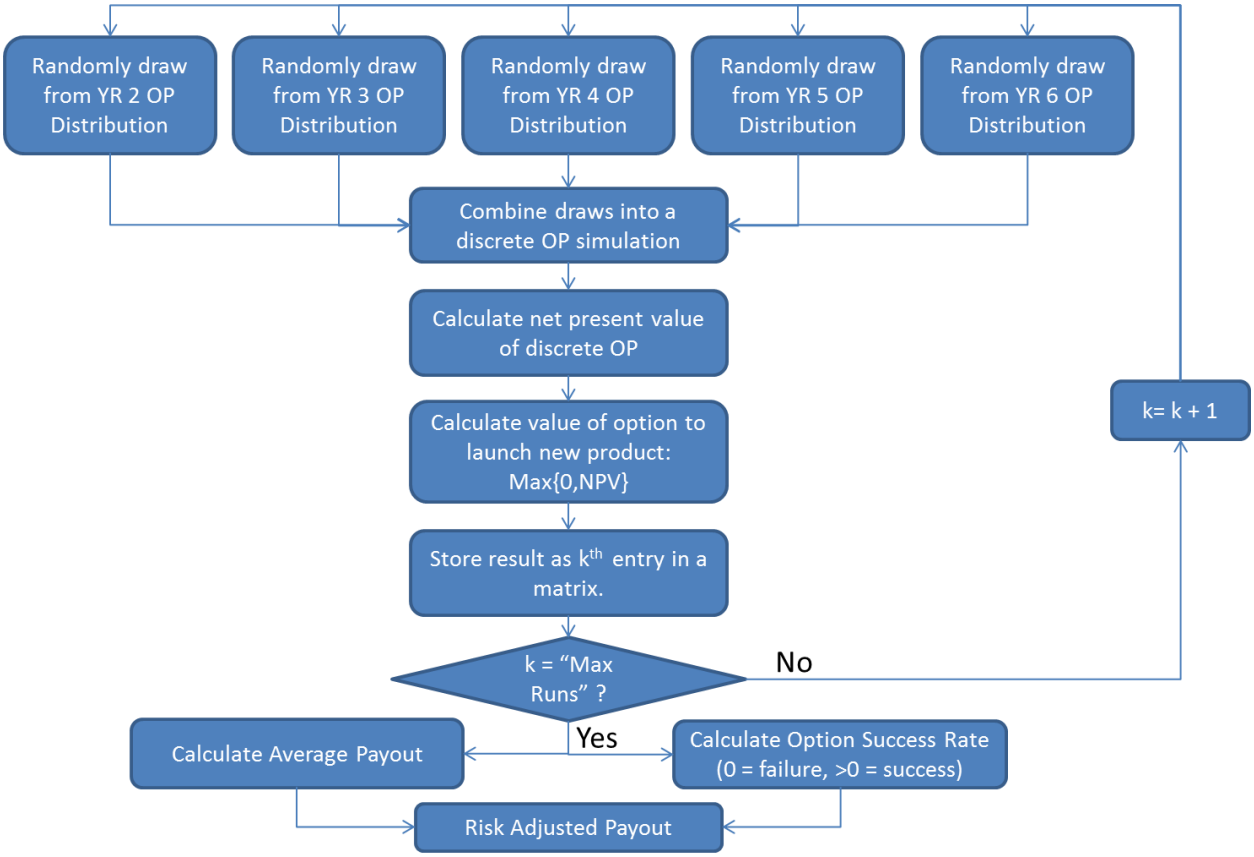


Figure 224: Flowchart of the simulation process

The value of the option for each run in the simulation is averaged to be $394000 in our example. A number of the runs result in a negative NPV, for which case the real option is valued at zero (see Figure 223b). Runs with a negative NPV indicate that the project would have been terminated, and capture the option to make contingency decisions based on the results of the R&D and the outlook for future sales. The risk adjusted payout is calculated with the related success rate.

$$\text{Risk Adjusted Payout} = (\text{successes per total runs})*(\text{mean RO value}) - (\text{R\&D cost}) \quad (310)$$

This establishes a single number that quantifies all of the uncertainty in the future to describe the value of investing in R&D. Using the operating profit assumptions in Figure 218,

the success rate is about 70% and the risk adjusted payout for investing in R&D is $182,000. For the values used in this example, the option to invest in maturing the technology to enable PHM for electronics is a profitable option.

The real option value distributions from this method are not lognormal, like in a Black-Scholes analysis, but rather generic to the details of the profit simulation. This is an advantage because the Black-Scholes approach is constrained to only modeling the future with a restrictive set of assumptions. As long as the Monte Carlo simulation converges, any method and any distribution can be used to model future operating profits. Therefore the presented method can be tailored to meet the specific needs of a wide variety of practical situations.

14.7 Comparison of Real Option Analysis with Net Present Value Approach

The presented approach is an extension of traditional NPV analysis. Using a traditional approach and the cash flow predictions for the most probable scenario, the net present value for the project can be calculated at $164347. Compared to the value derived from the real option approach, the NPV calculation undervalues the option to build a product with PHM capability. More importantly the traditional NPV analysis provides no insight into the risk associated with the investment or the option to abandon if the future is not as predicted.

$$NPV = \frac{-95000}{(1+0.15)^0} + \frac{-1x10^6}{(1+0.15)^1} + \frac{116195}{(1+0.15)^2} + \frac{275374}{(1+0.15)^3} + \frac{406680}{(1+0.15)^4} + \frac{557682}{(1+0.15)^5} + \frac{731335}{(1+0.15)^6} + = \$164347 \tag{311}$$

14.8 Comparison with Decision Tree Modeling Approach

Another alternative to real option analysis is the use of a decision tree. Decision trees are known to over value an option [Mathews 2007]. In our example, the decision to launch a product with PHM capability is made in year one. Using the operating profit distributions from Figure 218, the present value of the investment at year one, or the decision year can be calculated. Based on this simulation, the project would be profitable 77% of the time, the remainder of the time the project would be cancelled. The mean value of the profitable portion of the distribution is about $684000. Plugging this information into a decision tree, Figure 226, it can be shown that the decision tree over values the option to launch a product enabled with PHM technology.

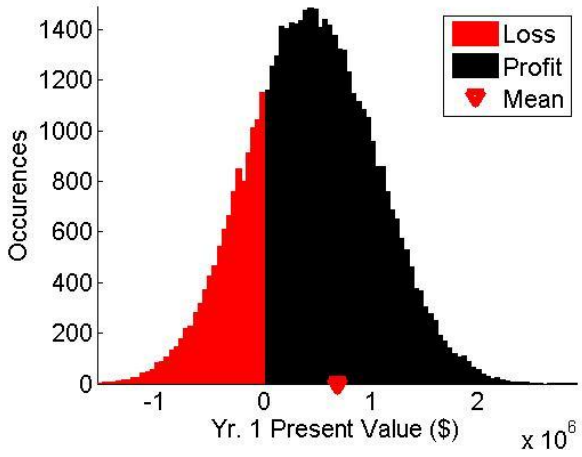


Figure 225: Present value of profit predictions at the decision point (year 1)

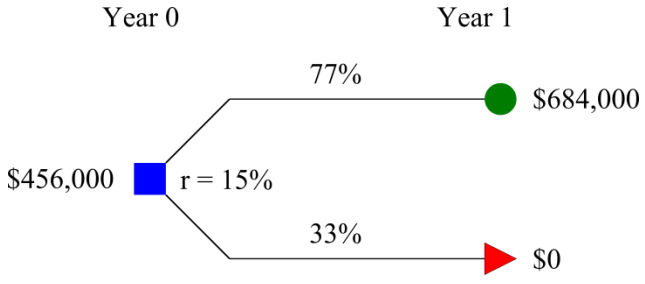


Figure 226: The decision tree approach can be shown to over value the option

14.9 Convergence of Monte Carlo Simulations

Convergence of the Monte Carlo simulation will vary depending on the specifics of your simulation and the number of parameters with distributions. Convergence needs to be proven before making any decisions based on the results. To test convergence, simulations were conducted with a varying number of runs per simulation. Thirty independent simulations were performed at each level, and the output prediction of the option value was recorded. Means and averages were calculated for each group of 30 simulations.

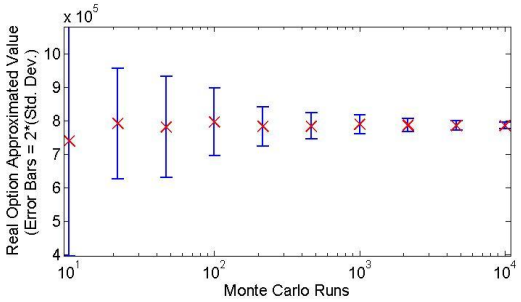


Figure 227: Convergence of the Monte Carlo Simulation

The spread of the uncertainty associated with the simulation decrease approximately as $\frac{1}{\sqrt{M}}$ where M is the number of runs per Monte Carlo simulation.

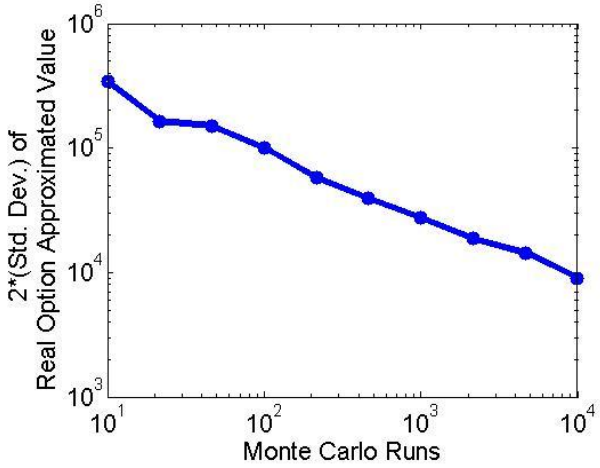


Figure 228: Convergence of uncertainty in the Monte Carlo simulation

14.10 Conclusion

A defendable method for assessing the cost of investing to mature PHM technology for electronics has been presented. To include the technical and financial uncertainty involved in assessing the cost of a PHM implementation, a real option approach has been used. Due to practical limitations with applying the Black-Scholes equation, an alternative called the Datar-Mathews method has been used to assess the value in investing in the R&D project for prognostics. The DM method is an extension of a DCF analysis, which most engineers and managers are familiar with. The advantages over a standard DCF approach include capturing uncertainty in the future and including the option of canceling the project as part of the value of the option. Alternatives to the real option analysis, NPV and decision trees, have been shown to under and overestimate respectively the value of the option to build PHM enabled circuit boards. The presented techniques are not limited to a single modeling methodology or a particular usage scenario, so they are useful in developing cost assessments for a wide range of applications.

14.11 Appendix

14.11.1 Derivation of operating profits for each year for three different scenarios

Relying on a single profit estimate over an extended period of time is risky when calculating the expected value of a project. A high/medium/low approach is used to allow for uncertainty in the prediction process. It is argued that anything more sophisticated than this type of estimation may put too much faith in a difficult and error prone prediction process, but that it is superior to a single profit prediction. There are however no restrictions on how operating profit estimates are obtained.

14.11.2 Unit Cost

The unit cost for a single board was calculated as the composite of the cost of the PCB, functional circuitry, and additional circuitry needed for PHM. Furthermore a premium was added for the additional real estate that the PHM circuitry will take up on the board. It is assumed that adding PHM circuitry to the board will cause design choices to be made that will increase the cost of producing the circuit board. For example smaller traces may be needed to properly route the functional components in a smaller area, increasing the cost of the board. It is assumed that eight channels of PHM monitoring will be included on each board. This would allow the monitoring of eight corner solder joints. This design decision is influenced by the limitations of the SCANSTA476 voltage monitor which has eight input channels. Details for the three prediction levels are shown in Figure 231.

The quantities column indicates that for example a resistor may be priced at $0.087, but 24 resistors are required per board. Prices are shown on a per component basis, but the extended price accounts for the full quantity required per board. PCB board costs are derived from low production run (<100) of off shore PCB fabrication costs for a 7” by 7” test board approximately the same size as JDES test board 3. As described in a later appendix section, only a few capacitors and resistors are needed to build the PHM monitoring circuitry. To extract the PHM signals from the board through existing diagnostic channels a voltage monitoring chip, the SCANSTA476, is required. All passive component pricing is representative of the cost at price breaks for <100, 100 and 1000 piece order from Newark.com. Labor and overhead were estimated for a high volume production line that is running low volume runs. This represents a worst case scenario where a significant amount of startup time is required to produce relatively few PHM capable boards.

	Pessimistic	Most Probable	Optimistic	Qty
4 layer PCB	25	20	15	1
Functional Components	100	75	50	1
*PHM Circuit Components				
Resistors	0.087	0.054	0.021	24
Capacitors	0.055	0.052	0.042	16
Scansta476	7.1	6.4	5.04	1
Labor per board	100	50	25	1
Overhead per board	100	50	25	1
Real estate premium	1.2	1.1	1.05	
Extended Cost	$ 402.08	$ 223.88	$ 127.28	

*Assumes eight monitoring channels

Figure 229: Three possible levels of unit costs for a PHM enabled circuit board

14.11.3 Sales Price

PHM capable circuit boards are a new product that has no historical records for pricing. Therefore pricing was based on a percentage markup. Based on the high level discussion of justifying the added cost of including PHM capable circuit boards for situations where safety, cost, and logistics were important to an organization, the markup for PHM capable circuit boards could vary widely. In safety related applications the markup could be quite high, but in cost sensitive maintenance applications the markup would have to competitive with traditional maintenance costs. Based on these criteria three markup levels were chosen to cover a wide range of possibilities.

Pessimistic	Most Probable	Optimistic
1.50	1.85	2.20

Figure 230: Pricing markup based on the uncertainty of the customers willingness to risk unplanned failures

Pessimistic	Most Probable	Optimistic
$ 603.12	$ 414.18	$ 280.01

Figure 231: Unit costs for the three scenarios

14.11.4 Annual Sales Quantity

Predicting anticipated sales for a product that does not exist in any form today is challenging. Three levels of sales estimated to match with three possible situations. In the pessimistic scenario the need for PHM enabled products remains niche, and only a few units are sold each year for use in relatively expensive applications. The most likely scenario is the product in accepted in a slightly higher volume product, but still faces adoption challenges in the first year. It is expected as a larger database of averted failures are documented, and longer running cost comparisons are completed the demand for PHM related products will increase drastically. For example, the Toyota Prius was a good example of untested technology that users were slow to adopt. To sell the first one million Prius' took 11 years, while the next million were sold in two years. Even the optimistic scenario does not assume widespread early adoption, but bases sales numbers on the assumption that an external government contract would increase sales volumes.

	Pessimistic	Most Probable	Optimistic
Initial Units Produced	2000	4000	8000
Annual Sales Growth	1.1	1.15	1.2

Figure 232: Possible levels of unit sales quantities

Sales growths were based on generic assumptions about usual growth rates for products in the pre-adoption stage.

14.11.5 Recurring Costs

Support for a new product will require engineering support. For the sake of argument a new team of between 4 and 6 engineers will be required to support a new PHM capable circuit board. A portion of the engineers would be expected to interface directly with customers and

possibly provide onsite training. It is assumed that the engineers will be trained at the graduate level to handle the highly cross disciplinary nature of the new product.

	Pessimistic	Most Probable	Optimistic
Engineers (M. Sc.)	4	3	2
Annual Salary	$ 125,000	$ 100,000	$ 75,000
Benefit Costs	$ 125,000	$ 100,000	$ 75,000
Extended Cost	$1,000,000	$ 600,000	$ 300,000

Figure 233: Estimated levels of engineering support for new PHM product

14.11.6 Operating Profit

The compilation of the profits and costs described above are shown in the table below. Redesigning a circuit board is an expensive undertaking and the launch cost of the project is estimated to be very high at $1M. The launch cost would encompass all of the electrical engineering design and layout required to add the PHM circuitry to the board. The launch cost would also include all tooling, product qualifications and other costs associated with designing a new board. Compared to the single time product launch cost, the R&D expenditure is very insignificant. The pessimistic/most probable/optimistic approach allows the incorporation of uncertainty into the calculation. Using the three levels of operating profits, a risk adjusted payout can be calculated. With the calculated risk adjustment decision makers can quantify the probability of the R&D investment returning a good payout.

Year	0	1	2	3	4	5	6
Pessimistic							
Unit Sales			2000	2200	2420	2662	2928
Unit Cost			$ 402	$ 402	$ 402	$ 402	$ 402
Unit Price			$ 603	$ 603	$ 603	$ 603	$ 603
Revenue			$ 402,082	$ 442,290	$ 486,519	$ 535,171	$ 588,688
Recuring Costs			$(1,000,000)	$(1,000,000)	$(1,000,000)	$(1,000,000)	$(1,000,000)
Launch Cost		$(1,000,000)	$ -	$ -	$ -	$ -	$ -
R&D Cost	$(95,000)	$ -	$ -	$ -	$ -	$ -	$ -
Operating Profit	$(95,000)	$(1,000,000)	$ (597,918)	$ (557,710)	$ (513,481)	$ (464,829)	$ (411,312)

Year	0	1	2	3	4	5	6
Most Probable							
Unit Sales			4000	4600	5290	6084	6996
Unit Cost			$ 224	$ 224	$ 224	$ 224	$ 224
Unit Price			$ 414	$ 414	$ 414	$ 414	$ 414
Revenue			$ 761,195	$ 875,374	$ 1,006,680	$ 1,157,682	$ 1,331,334
Recuring Costs			$ (600,000)	$ (600,000)	$ (600,000)	$ (600,000)	$ (600,000)
Launch Cost		$ (1,000,000)	$ -	$ -	$ -	$ -	$ 1
R&D Cost	$ (95,000)	$ -	$ -	$ -	$ -	$ -	$ -
Operating Profit	$ (95,000)	$ (1,000,000)	$ 161,195	$ 275,374	$ 406,680	$ 557,682	$ 731,335

Year	0	1	2	3	4	5	6
Optimistic							
Unit Sales			8000	9600	11520	13824	16589
Unit Cost			$ 127	$ 127	$ 127	$ 127	$ 127
Unit Price			$ 280	$ 280	$ 280	$ 280	$ 280
Revenue			$ 1,221,857	$ 1,466,229	$ 1,759,474	$ 2,111,369	$ 2,533,643
Recuring Costs			$ (300,000)	$ (300,000)	$ (300,000)	$ (300,000)	$ (300,000)
Launch Cost		$ (1,000,000)	$ -	$ -	$ -	$ -	$ 1
R&D Cost	$ (95,000)	$ -	$ -	$ -	$ -	$ -	$ -
Operating Profit	$ (95,000)	$ (1,000,000)	$ 921,857	$ 1,166,229	$ 1,459,474	$ 1,811,369	$ 2,233,644

Figure 234: Three levels of predicted operating profits. A summary of this chart is presented as Figure 216 in the main body of the text

15 PHM for the Micro-Coil Spring Interconnect

In an effort to meet reliability requirements for long term human presence in space without the need for resupply, a new interconnect for grid array packages has been developed. The interconnect utilizes beryllium copper springs which are 0.05 inches in height as interconnects between the package and PCB. These novel interconnects are known as micro coil springs (MCS). The configuration is approximately the same height as copper column interconnects, but has increased compliance compared to traditional column interconnects. Because the interconnect is still in the design stage, the feasibility of integrating prognostic health management capability into the interconnect is being studied. Failure prognostics, or the prediction of impending failure for individual components, would help ensure the reliability of systems deployed on long duration space missions and provide warnings of potential failure with adequate time to formulate contingency plans. Prognostic monitoring circuitry, prediction algorithms, and performance validation are discussed for micro coil packages subjected to JDEC standard drop testing.

15.1 Introduction

The correct operation of electrical subsystems are critical for manned space flight. New micro coil spring interconnects developed at the Marshall Space Flight Center [Strickland 2011], are a proposed solution for creating increased confidence in critical electrical subsystems for extended human presence in space without the need for resupply. As part of a larger testing program, a small subset of boards has been proportioned for a preliminary evaluation of drop and shock reliability. Because the interconnect is still in its design phase, the feasibility of monitoring the interconnects for wear, and then predicting the remaining useful life of the component from a

live sensor data stream is being investigated. If prognostic capability can be integrated into a circuit board during the electrical design phase, significant improvements in reliability may be realized. Incorporating prognostic capability into a circuit board after design and fabrication is very challenging. The authors are not aware of any techniques that allow real time interrogation of damage due to drop and shock loadings for in service electronics, without modifications specifically made to facilitate prognostics. Monitoring a system for wear, and predicting end of useable life is commonly known as prognostic health management (PHM) and allows the life cycle of the component being monitored to be actively managed as opposed to the reactive nature of responding to a failure diagnosis.

Previously a variety of grid array interconnects have been studied and shown to be good candidates for failure prognostics [Lall 2009a-d]. A wide variety of leading indicators of failure in electronics have been studied for fault diagnosis, fault mode detection, and failure prognostics [Lall 2004a-d, 2005a-b, 2006a-f, 2007a-e, 2008a-f, 2009a-d, 2010a-h]. Methods applicable for the real time sequential processing of monitored damage indicators have been demonstrated for BGA interconnects [Lall 2010a, 2011a,b]. Monitoring and failure prognosis for the micro coil springs is new.

In this paper a method of monitoring and prognosticating failure for micro coil springs (MCS) is demonstrated. Details of the JDEC standard drop testing and analysis of the MCS using high speed video and digital image correlation to extract full field strains during testing are discussed. Due to the increased compliance of the MCS, the interconnect experiences a noticeable relative displacement between the printed circuit board and the grid array component. This relative displacement has been observed and quantified. The required circuitry and its calibration needed to non-destructively monitor MCS for damage in a real time manner is

discussed. Hybrid prognostic algorithms influenced by both a preliminary understanding of failure mechanisms and data driven methods are created specifically for the MCS. Validation of the prediction algorithm on test cases and validation of the algorithms performance is quantified to benchmark the technique. A method for detecting algorithm divergence is shown to correctly detect erroneous results. Finally the PHM techniques demonstrated in the paper are compared against the traditional Weibull analysis to quantify the relative skill of the PHM predictions.

15.2 Test Vehicle

The novel geometry of the test vehicle is shown in Figure 235, and a schematic of the design with dimensions is shown in Figure 236. The assembly process for the MCS involves placing a package into a graphite fixture and vapor phase soldering the springs to the grid array component. A second graphite fixture is used to hold the populated package for assembly to the circuit board. Full details of this process are included in [Strickland 2011]. Each test board included a single MCS and a variety of traditional components for reliability benchmarking purposes. A single MCS was tested on each board to ensure that the stress profile seen by each component was as consistent as possible. Based on strain measurements discussed in the next section, strain histories at different locations on a JDEC style drop test board can vary widely. Table 24 shows the pertinent details of the MCS design. The MCS packages were daisy chained through all 400 interconnects.

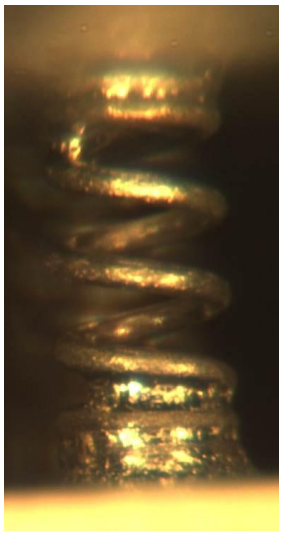


Figure 235: View of the micro-coil spring under an optical microscope at 20x magnification

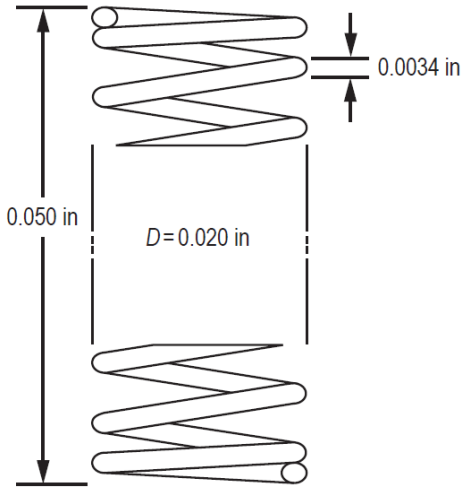


Figure 236: Schematic representation of the micro-coil spring

The electrical resistance of the daisy chained MCS interconnect is significantly higher than traditional BGA or copper column array interconnects. After assembly the MCS has a resistance of approximately 15 ohms. For comparison, daisy chained BGA components with a similar I/O count have a nominal daisy chain resistance on the order of a few ohms.

Table 24: Micro-coil spring package details

Parameter	Value
Interconnect	Micro-coil spring
Package size	22 mm
Package Type	Ceramic with aluminum heat sink
Package Alloy	Sn63-Pb37
I/O count	400
Grid Pattern	Full

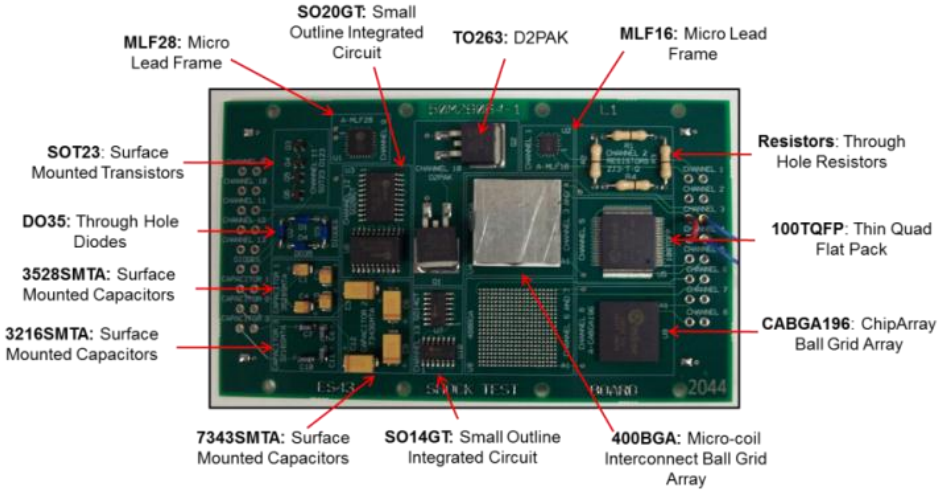


Figure 237: Micro-coil spring test vehicle

15.3 Experimental Setup

For the drop test matrix five MCS boards were tested according to JDEC standard JESD22-B111 drop tests [JDEC 2003]. The test was performed on a Lansmont drop test tower and controlled through a Labview graphical user interface. For each drop a time history of the impact pulse was recorded and monitored for compliance with the standard. The JDEC standard specifies that shock impulses remain within +/-10% of the stated 1500g, 0.5ms pulse. In practice tolerances of +/- 4% were obtained.

In addition to the impact pulse, a 20ms time history for each electrical continuity channels was recorded for each drop. A Lecroy WR-Xi digital signal oscilloscope was used to capture the shock impulse and continuity. LabView facilitated the transfer of the recorded time

history through the GPIB (IEEE-488) communication standard. The specifics of the circuit used to monitor the MCS are discussed in the health monitoring circuit section.

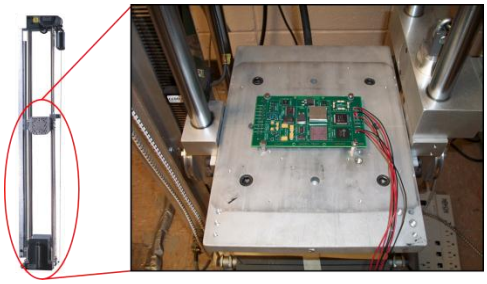


Figure 238: Representative drop test setup. For the data shown in this document the package was mounted face down on the drop table. The configuration shown in the picture allows the individual MCS to be monitored with high speed video during a drop event.

Because the interconnect is new, and failure mechanisms are still being studied, a variety of techniques were used to quantify the dynamic response of the system when subject to drop loadings. First a traditional strain gage was mounted on the board to acquire baseline data using a Vishay 2311 signal conditioning amplifier. Next the board was monitored during drop testing with high speed digital video cameras. Digital image correlation was used to extract full filed strain values from the high speed video.

Using the high speed cameras and a zoom lens focused only on the MCS reveals a relatively large amount of relative motion between the package and the circuit board. Figure 239 quantifies the relative displacement observed between the package and PCB. The influence of this dynamic behavior during drop testing on failure modes is still being studied. As a result of this uncertainty in the failure modes a hybrid model/data driven approach was utilized for implementing prognostics.

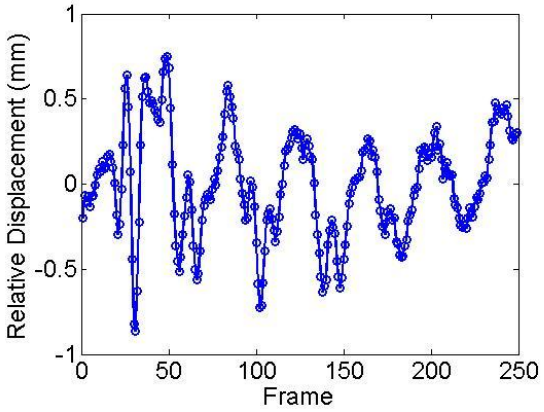


Figure 239: Relative displacement of package vs circuit board with MCS interconnects

15.4 Health Monitoring Circuit

The health of the MCS was interrogated non-destructively using electrical resistance as an easily measured quantity that correlates with damage. As the MCS accumulates damage in the form of plastic deformations and partial cracks, the geometry of the interconnect changes. Changes in geometry of the interconnect causes changes in the path through which current flows, and finally is measured as a change in electrical resistance of the interconnect. Previously for BGA components very precise resistance measurements [Lall 2009a-d] were utilized to monitor interconnects for damage. A special technique called resistance spectroscopy was required to measure the small changes in resistance in the BGA interconnects. The increased initial starting resistance of the MCS allowed a simpler measurement technique to be implemented.

Typical BGA components have a daisy chain resistance on the order of a few ohms. When BGA packages become damaged they experience changes in resistance on the order of a micro-ohm, followed shortly thereafter by a full interruption of electrical continuity (infinite resistance), which is the traditional definition of failure. A robust measurement technique for monitoring small changes in resistance is to use a Wheatstone bridge. The Wheatstone bridge

circuit is used extensively with resistive strain gages. Variations of the circuit have been previously used for prognostics [Lall 2010a, 2011a,b]. To use the Wheatstone bridge the daisy chain resistance of the package being monitored is integrated as one of the resistive arms of the bridge. As the package becomes damaged, the subsequent changes in electrical resistance can be detected with the Wheatstone bridge. Other methods exist for making resistance measurements including micro-ohm meters and four point kelvin setups, but the Wheatstone bridge has four practical characteristics which make it easier to use. Every package has a slightly different resistance. The Wheatstone bridge measures changes in resistance, not absolute resistances, therefore small manufacturing variations do not need to be calibrated out and do not affect the PHM algorithms. Electronics routinely experience wide fluctuations in operating temperatures. Unlike ceramic resistors, metal conductors have a high thermal coefficient of resistance. It is easy to temperature compensate the Wheatstone bridge by picking appropriate resistors. Wheatstone bridges are very simple, comprising of only a few resistors. Real estate on the surface of a PCB is always at a premium, so the implementation cost of adding a Wheatstone bridge circuit to a PCB to facilitate embedded PHM capability would keep additional costs at a minimum. Finally the technique can be adapted to monitor any number of interconnects. For example the techniques demonstrated in this paper can be modified to monitor only the corner solder joints of a fully functional package[Voutilainen 2009, 2010].

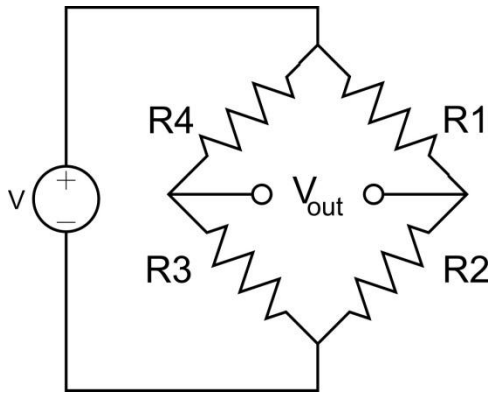


Figure 240: Schematic of a Wheatstone bridge circuit

The traditional Wheatstone bridge output equation that relates the resistances in the arms of the circuit to the voltage measured across the bridge (Figure 240: Schematic of a Wheatstone bridge circuit) is stated as [Dally 2005 pg. 324]

$$V_o = \frac{R_1 R_3 - R_2 R_4}{(R_1 + R_2)(R_3 + R_4)} V \tag{312}$$

The driving voltage, V, is a constant voltage source and is not affected by changes in the resistance of the bridge. Assuming that the bridge is initially balanced ($R_1 R_3 = R_2 R_4$) the output equation can be rewritten as

$$\Delta V_o = \frac{r}{(1+r)^2}\left(\frac{\Delta R_1}{R_1} - \frac{\Delta R_2}{R_2} + \frac{\Delta R_3}{R_3} - \frac{\Delta R_4}{R_4}\right) \quad where, r = \frac{R_2}{R_1} \tag{313}$$

Note that higher order terms have been neglected in this formulation and that this particular formulation is only valid for small changes in resistance. When trying to measure very small signals such as changes in resistance caused by damage to solder interconnects, optimizing sensitivity is important. The first term in the equation, $r/(1+r)^2$,scales the output, and so maximizing this term maximizes sensitivity. To initially balance the bridge and obtain maximum sensitivity all resistors should be equal.

A BGA package with a nominal daisy chain resistance of two ohms, when incorporated into a bridge with two ohm resistors in each arm it has a total equivalent resistance of two ohms. If the bridge is driven with a constant 5V source, then the circuit draws 2.5 Amps and tries to dissipate 12.5 Watts, which would surely melt/destroy the resistors in a matter of minutes. To keep currents at acceptable limits, 500 ohm resistors are used in each arm of the bridge. The 500 ohm resistor in series with the resistance of the package, shown in Figure 241, unfortunately decreases the sensitivity of the bridge by making the $\frac{\Delta R_x}{R_2+R_x}$ term much smaller, where Rx is the resistance of the package being monitored. To balance the bridge resistor R1 is replaced with a potentiometer. The MCS has a resistance of approximately 15 ohms, and the interconnects change resistance upon incurring damage, but prior to failure of electrical continuity.

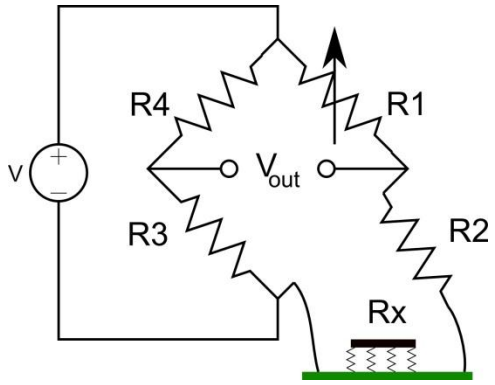


Figure 241: Wheatstone bridge with the resistance of the MCS daisy chain included as an arm in the bridge

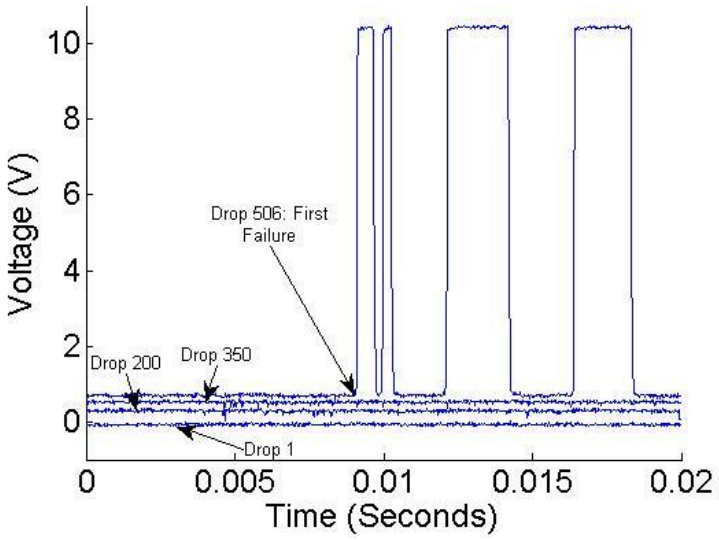


Figure 242: Raw data recorded from Wheatstone bridge for test number one.

Figure 242 shows the raw data that is recoded using a Lecroy WR-Xi digital signal oscilloscope at a frequency of $5x10^6$ samples/second. During drop 506, of test number one, the package experiences an intermittent open event and is defined as failed. The voltage signal is measured as the output of the Wheatstone bridge and therefore is proportional to change in resistance. The calibration curve in Figure 243 shows the close correlation between theory and observed results for the bridge. Above 20 ohms the circuit saturates and does not provide a reliable measurement.

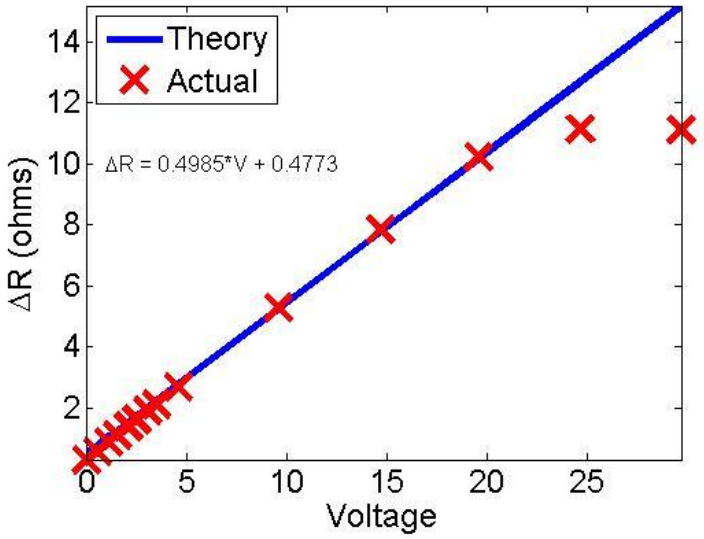


Figure 243: Calibration curve for Wheatstone bridge

15.5 Health Monitoring Algorithm

The Wheatstone bridge setup is as a non-destructive sensor for interrogating the health of the solder interconnects being monitored. To facilitate prognostic health management the raw stream of sensor data must first be converted into information appropriate for decision making. Figure 244 highlights the major sections of the health monitoring algorithm. Initial data processing is handled by a recursive filtering algorithm, the Kalman filter. The Kalman filter tracks and smooth's the noisy data stream coming from the sensor. Next the current state of the system is used as a starting point to prognosticate (predict) into the future when the component will fail. Lastly the prediction is converted into statistically defendable information appropriate for decision making.

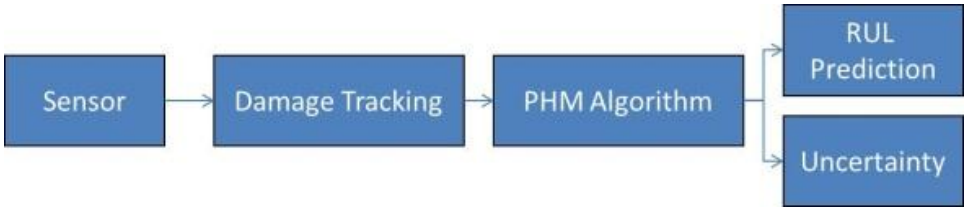


Figure 244: Process flow of information through the various parts of the PHM system

A variety of health monitoring algorithms have been utilized for implementing PHM on BGA components, including the Kalman filter (KF) implementation shown in this paper. All of the algorithms are recursive in nature, which is beneficial for two reasons. First, recursive algorithms process data in real time as a stream of information from a sensor. This is opposed to traditional curve fitting algorithms where all the data is processed in bulk. The task of measuring and storing in memory a full time history for every component being monitored becomes un-tractable in practice. Secondly filters efficiently reject noise from the measured signal. The algorithms demonstrated in this paper are generic enough to be applicable to any physical system that is wearing out assuming you have sufficient domain specific knowledge of failure mechanisms. The applicability of the KF to prognostics was demonstrated previously on BGA components, and is adapted for MCS in this paper.

To implement the KF an approximate damage model of the MCS must be created. At this stage in development of the MCS the system model is very coarse. As more information about specific failure mechanisms becomes available, improved models would be expected to enhance the PHM algorithms performance. The remainder of the paper will be dedicated to developing an approximate system model, verifying correct operation of the PHM algorithms, and validating algorithm performance.

15.6 Prognostics Hindecaste

This section will present a proof of concept for a prognostic system to monitor and manage the health of MCS interconnects using data acquired during run to failure drop tests. To borrow terminology from hurricane forecasting, a forecast is the real time prediction of future events based on observations. A hindecaste is a simulated real time prediction of future events, using previously observed data. New hurricane models are validated using the hindecaste method on more than one hundred years of climatology data [Rappaport 2009]. The prediction scheme demonstrated here uses a database of five run to failure tests. Leave one out cross validation will be used to quantify the performance of the algorithms given the relatively smaller database available for validation.

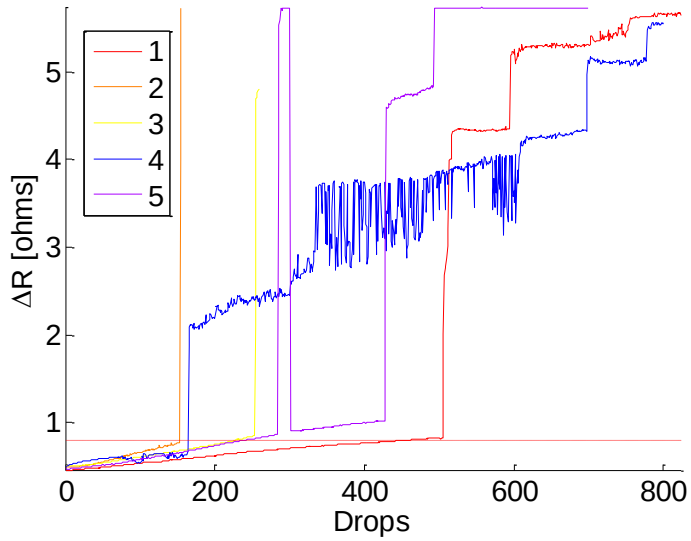


Figure 245: Time history of five run to failure drop tests. The dotted red line represents the average resistance increase of each component at failure

Figure 245 shows the time history of the MCS's resistance for five fun to failure tests. When the resistance increases dramatically, the electrical continuity has been broken and the part will soon have an infinite resistance and therefore is defined as failed. The challenge is to

sequentially process the data stream before the part fails and predict the remaining useful life based on the information available up to the current time.

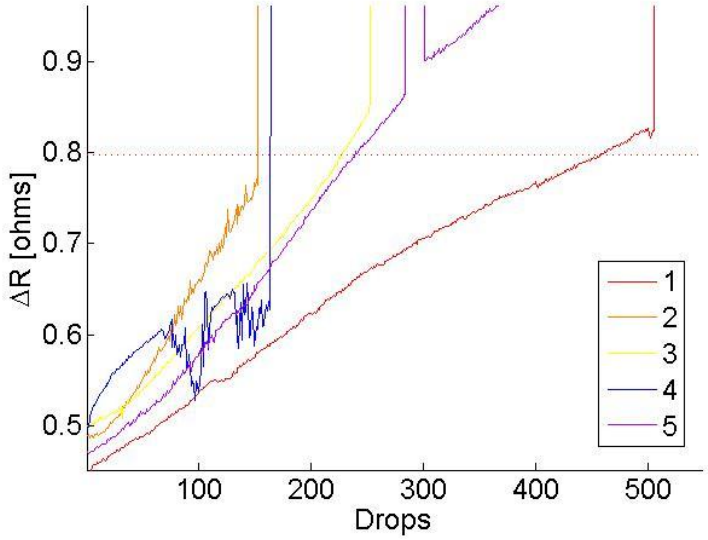


Figure 246: Zoomed view of the resistance time history plots. The dotted red line represents the average resistance increase of each component at failure

The experimental setup section described how for each drop a 20ms time history of the output from the Wheatstone bridge was recorded. The data shown in Figure 245 and Figure 246 represent the average resistance over each 20ms time history to reduce the dimensionality of the data. From visually inspecting the data it can be seen that during each test the component failed after an increase in resistance on average of about 0.8 ohms. Table 25 details the failure time and resistance of each test. The average resistance increase at failure, will be used as the failure threshold as described in the validation section.

Table 25: Drops to failure, and corresponding change in resistance

Run	Drops to Failure	Change in Resistance at failure (Ω)
1	506	0.824
2	154	0.771
3	254	0.850
4	166	0.677
5	285	0.863

The nearly linear increases in resistance up to failure was not expected to occur before testing began. Based on previous experience from BGA packages [Lall 2011a], the initiation of failure coincided with an increasingly noisy signal that trended exponentially to failure. Because the experimentally observed linear increase in resistance does not match with previous experience we dedicated significant effort to eliminate it from testing results. Initially it was assumed that the increase in resistance was a relic of the experimental setup. At first it was hypothesized that the slow increase in resistance was the result of Joule heating. Joule heating was eliminated as the cause of the resistance ramp up by leaving the board wired into the test circuit for eight hours without performing any drops. Over the course of the static test no ramp in resistance was observed. Next it was hypothesized that the resistance ramp up was the result of the test circuit. This possibility was eliminated by wiring a traditional BGA package into the circuit and performing a drop test. No ramp up was visible when using a BGA package, only the MCS exhibited this unusual behavior. Again it was hypothesized that some unknown effect from the MCS was causing heating during the test. This possibility was eliminated by pausing a MCS test halfway through the test and leaving the monitoring circuit on. Again no ramp in resistance was ever observed without stressing the package during drop testing. Another hypothesis was that the continuity wires leading from the package to the Wheatstone bridge were the cause of the resistance increase. A variety of methods for securing the continuity wires to the drop table were tested, using both solid and braided wires. By process of elimination we ultimately decided that the linear increase in resistance was a result of damage occurring in the MCS interconnects.

Based on observations of the run to failure data, it is assumed that a first order model (y = mx + b) is sufficient to model failure. Therefore, two system states are needed to predict future values of a MCS change in resistance. The change in resistance of the MCS and the first derivative of change in resistance with time. The notation for the discussion below was taken from the excellent book [Zarchan 2000] and more details about the KF as it pertains to prognostics for electronics can be found in [Lall 2010a]. The discussion here will describe the components needed to create a filter for the purposes of tracking damage in MCS interconnects. The system states are denoted as

$$x_k = \begin{Bmatrix} \Delta R \\ \Delta \dot{R} \end{Bmatrix} \tag{314}$$

Because only change in resistance can be directly monitored, the measurement equation is formulated as

$$z_k = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{Bmatrix} \Delta R \\ \Delta \dot{R} \end{Bmatrix} \tag{315}$$

By inspection the measurement matrix is $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$. The derivative of change in resistance with respect to time is assumed to be a constant value and can therefore be considered white noise denoted as $u_s \sim N(0, \Phi_s)$. A state space model of a system that only depends on its resistance and change in resistance is represented as

$$\begin{Bmatrix} \dot{\Delta R} \\ \ddot{\Delta R} \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} \Delta R \\ \Delta \dot{R} \end{Bmatrix} + \begin{Bmatrix} 0 \\ u_s \end{Bmatrix} \tag{316}$$

After this point, every KF implementation is the same, which makes the algorithms presented here applicable to any system as long as there is sufficient information about failure modes to write a state space model. By inspection the system dynamic matrix is

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \tag{317}$$

Subsequently the fundamental matrix can be shown to be approximately represented by a Taylor Series expansion

$$\Phi = e^{Ft} \approx I + Ft + \frac{(Ft)^2}{2!} + \cdots \tag{318}$$

Where I is the identity matrix. In the case being examined in this paper the second order terms and above of the fundamental matrix are zero, so only a first term approximation is meaningful.

$$\Phi = I + \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \cdots = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \tag{319}$$

Substituting the sampling time, Ts, for the variable 't' in the equation above provides the discrete form of the fundamental matrix

$$\Phi_k = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} \tag{320}$$

The continuous process noise matrix is influenced by our assumption that the change in resistance should be a constant

$$Q = \Phi_s \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{321}$$

The discrete process noise matrix is calculated from the continuous process noise matrix

$$Q_k = \int_0^{T_s} \Phi(\tau) Q \Phi(\tau)^T \, d\tau = \Phi_s \begin{bmatrix} \frac{T_s^3}{3} & \frac{T_s^2}{2} \\ \frac{T_s^2}{2} & T_s \end{bmatrix} \tag{322}$$

For each time step where a measurement becomes available the KF is used to track and estimate the state of the MCS.

Using the components derived above, the matrix Riccatti Equations [Zarchan 2000], are used to solve for the estimates of the resistance and ramp rate of the resistance.

1. Using the uncertain system model, predict where the next measurement is expected to be, $\bar{x}_k = \Phi_k \hat{x}_{k-1}$
2. Project the error covariance ahead, $M_k = \Phi_k P_k \Phi_k^T + Q_k$
3. Calculate the Kalman gain, $K_k = M_k H^T / (H M_k H^T + R)$
4. Make a measurement of the MCS change in resistance, z_k
5. Update the system state estimate, $\hat{x}_k = \bar{x}_k + K_k(z_k - H\bar{x}_k)$
6. Update the error covariance matrix $P_k = (I - K_k H) M_k$

The process noise term, Φ_S, and the measurement noise term, R, must be initialized at the start of the algorithm and their choice is discussed after the full tracking/prediction algorithm is presented.

The result of the KF implementation is a smooth estimate of both the change in resistance and the ramp rate of the change in resistance (Figure 247). Notice how the estimate of the ramp rate of change in resistance adjusts down as the slope decreases slightly in the last half of the test.

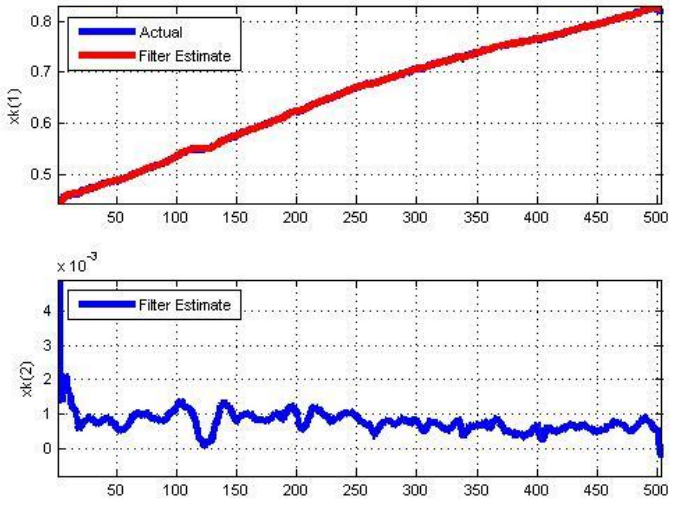


Figure 247: Filter results for the data set from test one. The top plot is the filter track of the measured resistance. The bottom plot is the estimate of the ramp rate of change in resistance (ie: derivative).

The assumptions outlined earlier assumes that the change in resistance can be modeled as a first order system. If the current state of the system is known, future states can be predicted by extrapolation.

$$\Delta R_{future} = \Delta R_{now} + (\dot{\Delta R})\Delta t \tag{323}$$

When the value of ΔR_{future} is taken as the failure threshold, and equation (323 can be solved for Δt, which is the equivalent to remaining useful life. Smoothed estimates of ΔR and $\dot{\Delta R}$ from the Kalman filter are used for the RUL calculation.

7. Report the remaining time until ΔR is estimated to break the failure threshold, $RUL_k = \frac{x_f - \hat{x}_k(1)}{\hat{x}_k(2)}$
8. Report the uncertainty in the estimate using [Swanson 2001], $\sigma_{RUL,k} = 1.86\left(\frac{P(1,1)}{P(2,2)}\right)$
9. Increment, $k = k + 1$, and return to step one

During the prediction step it is assumed that damage will accumulate at the same rate that has been observed previously. This is not ideal, but the future usage of the system cannot be deterministically known in advance. The prediction of the remaining useful life, and prediction uncertainty, for the interconnect at the current time is recorded for offline validation purposes, then the procedure begins again with a new measurement of the system. By using the KF, the ability to have an uncertain system model that adjusts to observations allows the algorithm to adapt to the case where the usage profile may be time evolving.

The KF is a recursive algorithm and must be initialized. The output from the Wheatstone bridge is near zero for a balanced bridge, so the resistance is initialized as zero ohms. No bridge can have a perfectly balanced starting condition, but the filter quickly converges to the proper estimate of the resistance. The resistance ramp rate has no physical justification for an initialization value, so an ignorant guess of zero is made. Again the filter quickly converges to a reasonable estimate for the resistance ramp rate. The process error covariance matrix effects the initial convergence of the filter and is set with arbitrarily high diagonal terms, which indicates that there is no confidence in the initialization conditions.

$$\hat{x}_0 = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{324}$$

$$P_0 = (1000)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{325}$$

Lastly the process noise term and the measurement noise term help the filter appropriately weigh the value of measurements when estimating system states. Based on standard deviations of steady state measurement when the test circuit was turned on, but no testing was occurring a value of 0.005 ohms was assigned to the measurement noise term. The standard deviation of the process noise term was taken as the expected variation in resistance

between time steps (0.005 ohms) divided by the expected number of time steps before failure (300 drops).

It is important to note the benefit of using a Kalman filter to estimate smooth derivatives. A numeric derivative would be calculated as

$$(\dot{\Delta R})_{numeric} = \frac{\Delta R_k - \Delta R_{k-1}}{\Delta t} \tag{326}$$

Taking a numeric derivative of noisy data results in white noise. A comparison between the numeric approach and the derivative estimated by the Kalman filter is shown in Figure 248. Smooth derivatives are essential for making accurate RUL predictions and motivate the use of the Kalman filter. Without a smooth derivative the RUL prediction would also be white noise.

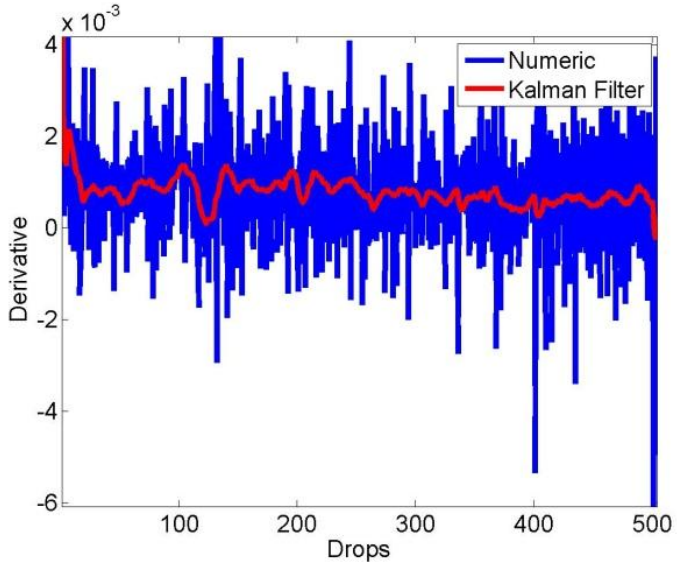


Figure 248: Comparison of the numeric derivative and estimated derivative for test number one. Smooth derivatives are essential for accurately predicting remaining useful life and motivate the use of the Kalman filter

Results of the prognostic hindecaste for test number one (Figure 245) are shown in Figure 247 (tracking) and Figure 249 (remaining useful life predictions). It can be seen that the filter closely tracked the measured data and estimated a nearly constant first derivative. A confidence

interval for the remaining useful life (RUL) prediction is reported with each prediction and is based on the KF's uncertainty in the state estimates. This helps decision makers quantify the confidence in the predictions coming from the PHM algorithm, and if the prediction is trust worthy. Results for test number one are representative of tests two through five. Results for the tests not shown in the body of the paper can be found in the appendix.

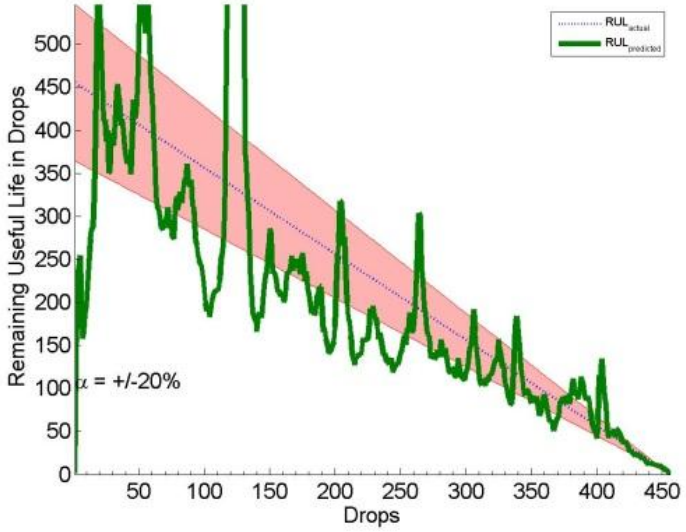


Figure 249: Remaining useful life predictions for the MCS

The next section will deal extensively with verifying and validating the operation of the PHM algorithm. Standard PHM metrics will be used to quantify the results, and a short discussion of each metric as it applies to the data set shown above will be discussed.

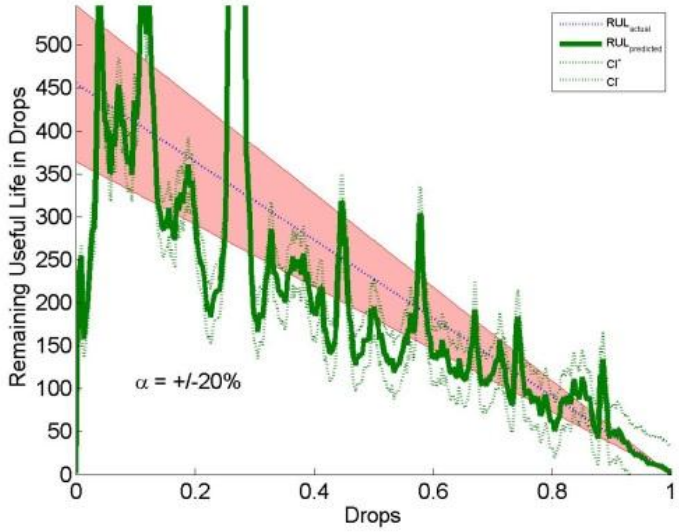


Figure 250: Alpha-lambda accuracy plot

The resulting performance function for the hidcaste on the first resistance history was 0.63. A perfect score would be zero, and the worst possible score is a value of one.

15.7 Verification and Validation

The methodologies of Verification and Validation (V&V), can be used to rigorously demonstrate the capability of a system to show that it was built correctly and performs as intended. V&V is a popular technique in system engineering and computational mechanics. It is easiest to understand V&V from a systems engineering standpoint, but the methodologies used in prognostics are probably closer related to computational mechanics since they rely heavily on algorithms and differential equations. From a systems engineering perspective, V&V can be loosely stated as

- Verification: Did you build your product correctly to the specifications of the technical drawing
- Validation: When you use your product, does it perform as intended

15.7.1 Verification

To verify the correct operation of the Matlab code used to perform the prognostic hindecasts a variety of different simulated test problems were used. Simulated test problems were used so that results could be compared against known ground truths. All implementations of the KF are identical beyond the formulation of the system model, which makes testing code easy. A variety of elementary dynamic systems are used to verify the correct operation of the code under a variety of circumstances. Test cases involving steady state signals, ramped signals, sinusoidal signals, and harmonically damped signals (Figure 251) were used to test the KF algorithm for correct performance.

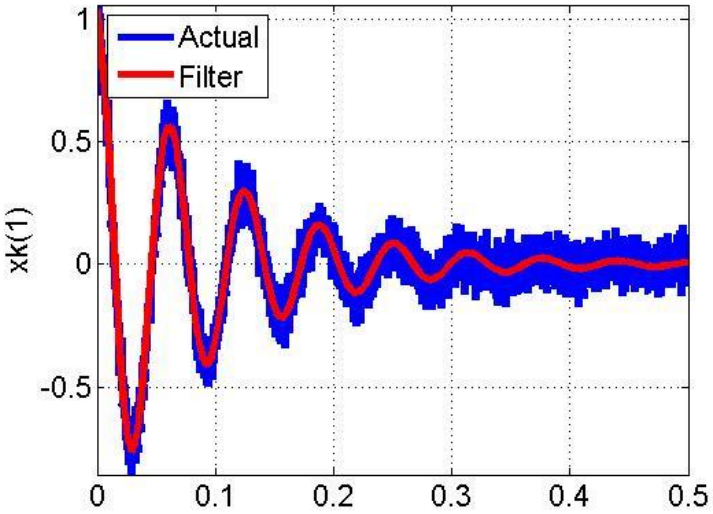


Figure 251: Verification of code using the case of a damped harmonic system

Health monitoring circuits were verified for correct operation by substituting a potentiometer in place of the daisy chain resistance of the package being tested as described in Figure 243.

15.7.2 Performance validation

Leave one out cross validation was used to validate the performance of the PHM algorithm using all five run to failure tests. In leave one out cross validation four of the tests are used to train the PHM algorithm, while the last test is used to validate the performance. In turn the process is repeated five times, leaving out a different test each time to validate. This procedure is alternatively called k-fold cross validation, where in this paper k=5. At the end of the validation an average performance score is reported.

Minimal training time is required for each validation run. The average voltage at failure is calculated using the four training sets. Then the PHM prediction algorithm is run on the "left out" data set as a hindcaste to validate the performance of the algorithm. After the conclusion of the hindcaste, the standard PHM metrics are used to calculate a cost metric that quantifies the algorithms performance in to a single number.

Training consists of calculating the average ΔR at time of failure to establish a failure threshold (x_f). The failure threshold defines failure when projecting system states into the future. For some of the validation runs this results in the situation where the component fails after the failure threshold is crossed. To handle this scenario if the ΔR estimated by the filter is above the failure threshold, RUL predictions are set to be zero, and the uncertainty in the prediction is taken as the standard deviation of the ΔR at failure for the training set. Using a failure threshold that does not perfectly correlate with actual failures adds error to the PHM predictions. For example the predictions previously shown for test number one show a systematic error when using the average failure threshold (Figure 252). Only after seeing all of the data can it be known that the failure threshold was not optimal. Leave one out cross validation simulates what the expected performance would be given the situation where the future is not known. Using the

failure threshold for run number one when testing run number one would constitute a logical fallacy known in data mining as a future information leak [Elder 2005]. Another example of a future information leak is taking a three day moving average, centered on the current day.

Table 26: Discrepancy between actual failure and failure threshold derived from training sets.

	Test Set	Mean Change in Resistance at Failure of Training Sets [ohms]	Actual Change in Resistance at Failure of Test Set [ohms]
Run #1	1	0.824	0.790
Run #2	2	0.771	0.804
Run #3	3	0.850	0.784
Run #4	4	0.677	0.827
Run #5	5	0.863	0.780

Table 27 summarizes the results for each run of the leave one out cross validation matrix.

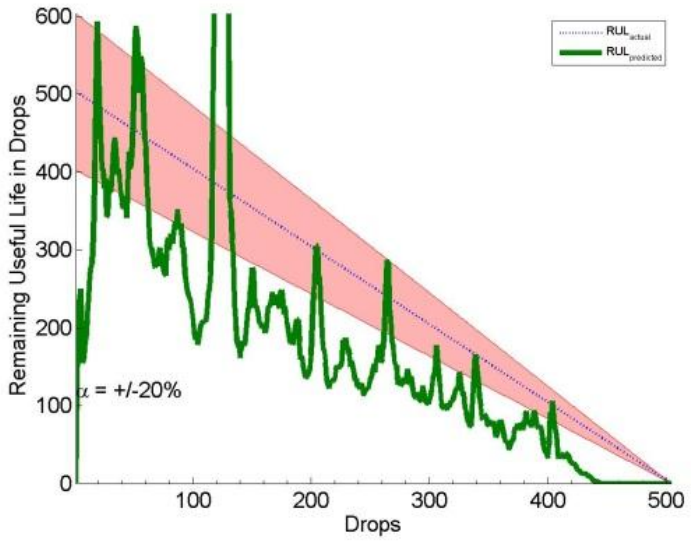


Figure 252: Systematic error introduced by using an average failure threshold derived from the training set.

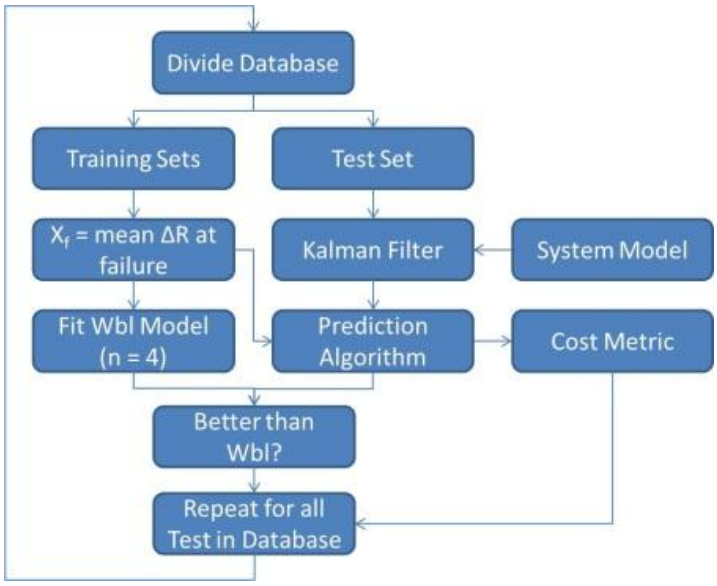


Figure 253: Process flow for leave one out cross validation discussed in this section and the comparison against traditional Weibull methods discussed in the next section

Table 26: Discrepancy between actual failure and failure threshold derived from training sets.

	Test Set	Mean Change in Resistance at Failure of Training Sets [ohms]	Actual Change in Resistance at Failure of Test Set [ohms]
Run #1	1	0.824	0.790
Run #2	2	0.771	0.804
Run #3	3	0.850	0.784
Run #4	4	0.677	0.827
Run #5	5	0.863	0.780

Table 27: Leave one out cross validation matrix. A cost metric of zero is a perfect score.

	Test 1	Test 2	Test 3	Test 4	Test 5	Cost Metric
Run 1	Test	Train	Train	Train	Train	0.6775
Run 2	Train	Test	Train	Train	Train	0.6857
Run 3	Train	Train	Test	Train	Train	0.533
Run 4	Train	Train	Train	Test	Train	0.8424
Run 5	Train	Train	Train	Train	Test	0.6738
					Average:	0.6825

15.8 Filter Divergence

Filter divergence is the term used when the filter does not accurately track the desired system states. If the system model is incorrect beyond the uncertainty provided for with the Φ_S term, then the residual between the estimated state and measured state increases. Theoretically if the filter is operating correctly the residual error should be bounded by the value of $\pm\sqrt{P_k(1,1)}$ for at least 68% of the measurements.

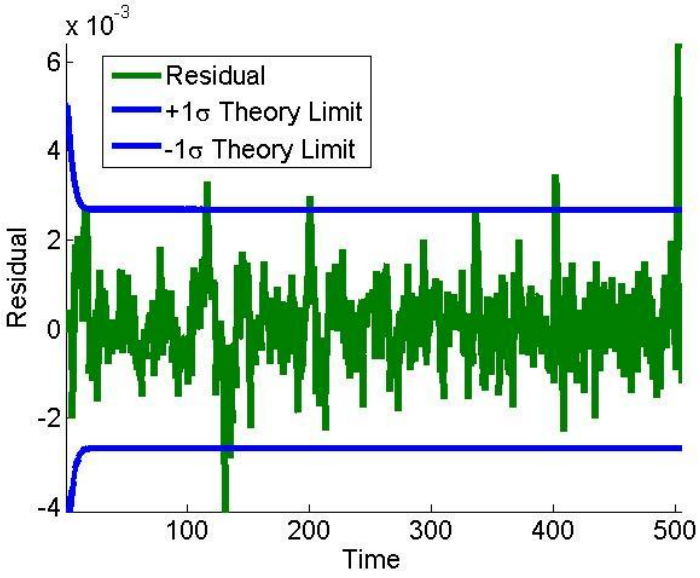


Figure 254: Residual between the estimated state and actual state demonstrates that the filter was tracking correctly for test number one

In the case of test number four it is hypothesized that a different failure mode occurred, and the resulting time history was significantly different than the other tests. The behavior of the system violated the assumptions of the system model and the results were not good. Monitoring for filter divergence would provide an online method of detecting the when reported RUL predictions may be in error.

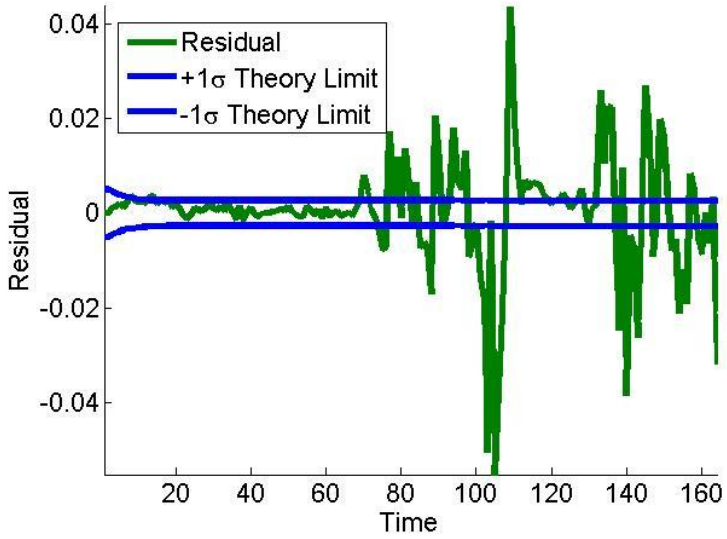


Figure 255: Residual between the estimate state and actual state for test #3. After drop 65 the filter divergence could have been used to warn that predicted RUL may be in error

15.9 Skill Against Weibull Baseline

In this section the skill of the PHM algorithm will be compared against a Weibull baseline to motivate the inclusion of PHM capability in the development of systems with MCS interconnects. This method is borrowed from weather forecasting [Roebber 1996]. Predicting the average failure time from Weibull analysis requires no skill on the part of the forecaster, assuming that they correctly operate the Weibull software. For this reason the failure predictions from a Weibull analysis are taken as an arbitrary baseline to compare the skill of the forecast generated with the PHM algorithms. Any baseline could be used, but historically Weibull

analysis has been widely used to model failure rates in electronics. The prediction algorithms in a PHM algorithm require a user to pick a recursive tracking algorithm, establish failure thresholds, choose a system model and a number of other inputs that effect the performance of the algorithm. If the resulting predictions from a PHM algorithm are better than predictions obtained from Weibull analysis then the prediction is said to be skill full. The skill of the forecast is quantified as

$$S_f = \left[\frac{e_b - e_f}{e_b}\right] * 100 \tag{327}$$

Where the skill of the forecast, S_f, is a function of the error between the baseline prediction of failure, and the actual failure (e_b), and the error between the forecast and the actual failure (e_f). For our example the skill of the PHM predictions, S_{PHM}, is a function of the baseline Weibull error and is denoted as e_{wbl}, the error from the PHM prediction denoted as e_{PHM}.

$$S_{PHM} = \left[\frac{e_{wbl} - e_{PHM}}{e_{wbl}}\right] * 100 \tag{328}$$

A skill greater than zero represents an improvement in forecasting accuracy over the baseline method. A skill less than zero represent a forecasting accuracy that is less effective than the baseline method. A positive skill score implies the forecast was skillful. Because the PHM algorithm outputs a time series of predictions the error was calculated as a mean squared error. Following the method outlined in Figure 253, during each leave one out run, the training set was used to calculate failure by traditional Weibull analysis. Then the predicted failure times from the PHM algorithm are compared against the mean failure time for the Weibull distribution as described in equation (328. The process is repeated for each leave one out run so that the relative skill for each test can be independently evaluated.

Figure 256 visualizes the improvement in the PHM predictions over the baseline Weibull approach for the data from the first test. At the end of the test the estimated ΔR is above the failure, but the part has not failed. When the estimate is above the failure threshold the algorithm predicts failure is imminent (RUL = 0) and reverts to using the standard deviation of the failure data to assign a level of uncertainty. Both the accuracy and precision of the PHM predictions are significantly better than the Weibull baseline for test one.

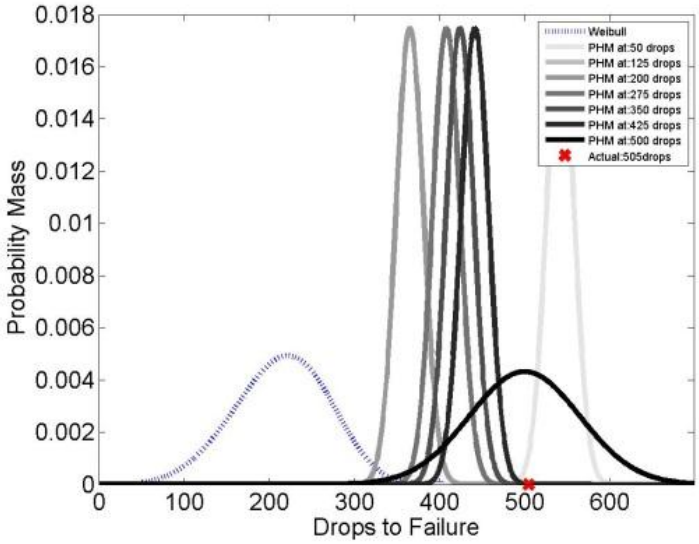


Figure 256: Overlay of a few PHM predictions on top of the baseline weibull distribution

15.10 Conclusion

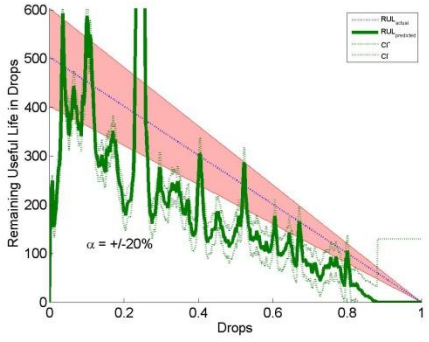
The feasibility of including prognostic health monitoring capability for a novel interconnect has been investigated. Correlations between changes in electrical resistance and failure have been used to prognosticate failure in components during JDEC standard drop testing. Details of the health monitoring circuitry were discussed. Leave one out cross validation has been used to estimate the performance of the failure prognostics techniques on data sets that do not exist yet. Validation and verification of the health monitoring circuitry, and failure prediction algorithms have been demonstrated. Filter divergence, a technique for online detection of

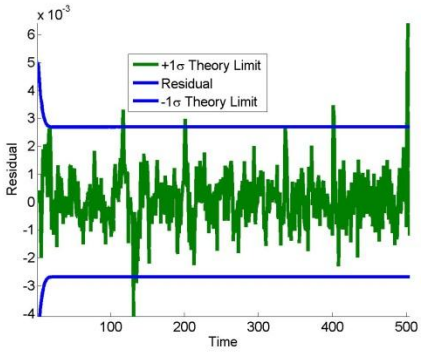
improper operation of the prediction algorithm, was shown to correctly identify the data set that reported bad predictions. Finally the results demonstrated in the paper were compared against traditional Weibull analysis to quantify the skill of the failure prognostics. An appendix including results from all of the run to failure tests was included to show the repeatability of the presented methods.

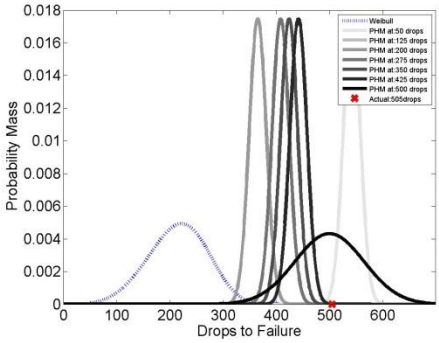
15.11 Appendix

The following is a summary of the results from the leave one out cross validation test. These results simulated the effectiveness of the PHM algorithms for data sets that been strictly divided into test sets and training sets. The results here represent the performance of the filter that could be expected in practice.

<u>Test Number One</u>
Drops to failure 506
ΔR at failure: 0.790 Ω
ΔR at failure based on tests {2-5}: 0.824 Ω
Skill against Weibull baseline: +38%

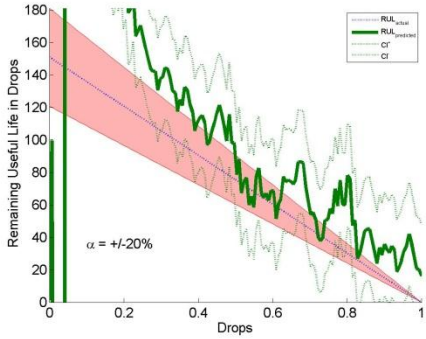


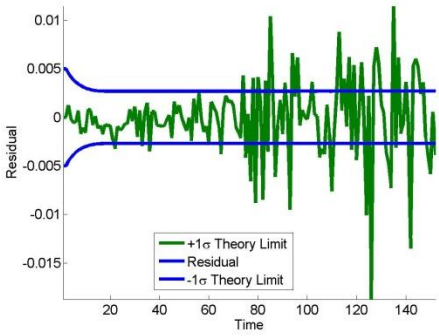


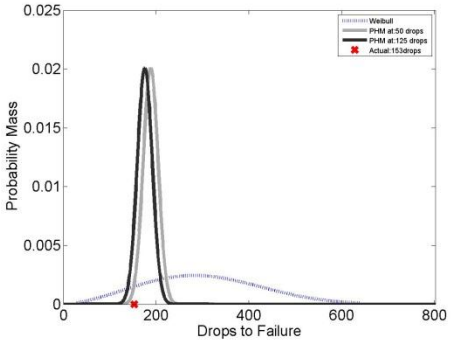


Comments on test number one: The tracking and performance of test number one was good. The systematic error in the PHM predictions is a result of the true failure being below the average failure threshold. This test represents an outlier in the data set and PHM successfully predicted significantly longer remaining useful life.

Test Number Two
Drops to failure 154
ΔR at failure: 0.804 Ω
ΔR at failure based on tests {1,3-5}: 0.824 Ω
Skill against Weibull baseline: +98%

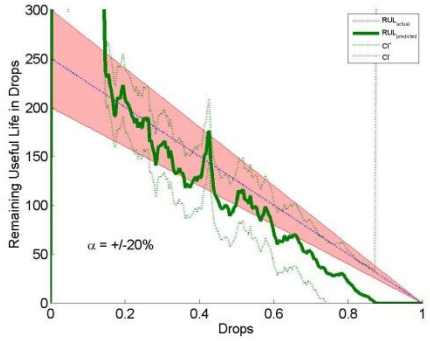


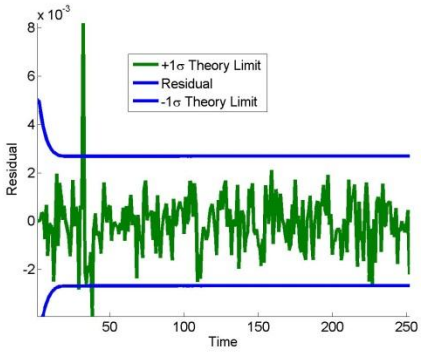


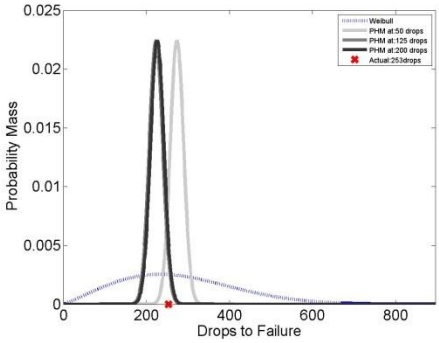


Comments: Tracking performance was moderate for test number two, but the algorithm did successfully provide advanced warning for a component that failed prior to the mean failure time. This test showed the best relative skill against the Weibull baseline.

Test Number Three
Drops to failure 254
ΔR at failure: 0.784 Ω
ΔR at failure based on tests {1,2,4,5}: 0.850 Ω
Skill against Weibull baseline: +6%

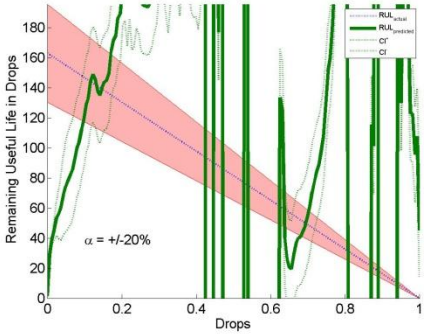


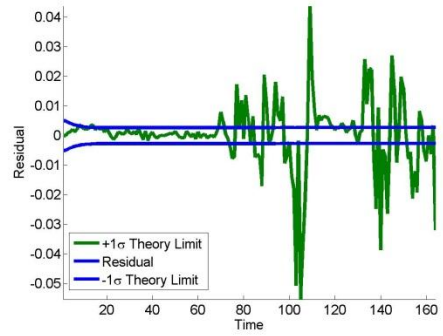


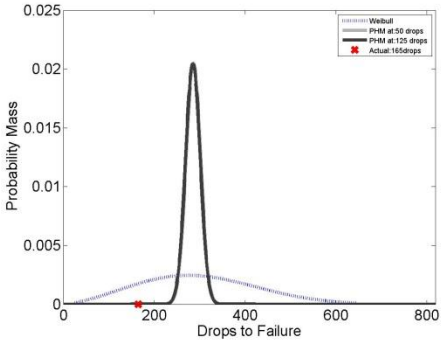


Comments: Tracking and prediction for this test were excellent, but the failure occurred very close to the mean failure time predicted by Weibull analysis so the relative skill does not appear as good as other tests. Note that the failure threshold was low for this test and the last 10% of the test the PHM algorithm thought failure was imminent.

Test Number Four
Drops to failure 166
ΔR at failure: 0.827 Ω
ΔR at failure based on tests {1-3,5}: 0.677 Ω
Skill against Weibull baseline: -4810%

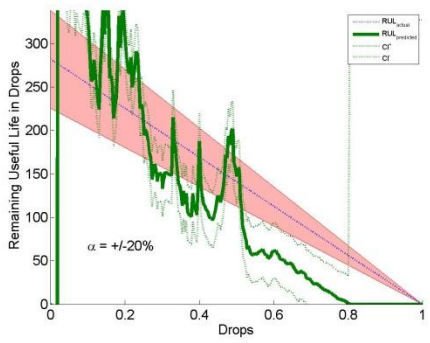


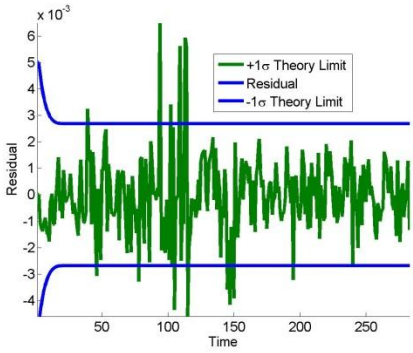


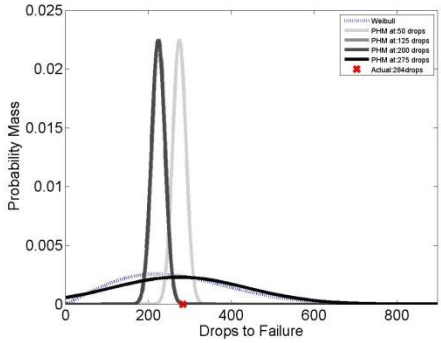


Comments: This test represents the bad apple from the test group. It is hypothesized that a failure mode different from the other four components occurred on this test. The test could have been easily substituted with a better test, but for the sake of academic integrity the first five consecutive test were used to form the failure database. Looking at the residual plot shows that the erroneous results could have been flagged as problematic in real time.

Test Number 5
Drops to failure 285
ΔR at failure: 0.780 Ω
ΔR at failure based on tests {1,3-5}: 0.863 Ω
Skill against Weibull baseline: -1429%







Comments: This test, like test number one, tracked and predicted well, but the discrepancy in the actual failure threshold and the average failure threshold resulted in a systematic error. The mean failure time according to Weibull analysis was only in error by twelve drops, so the relative skill of this prediction was not good. From looking at the overlay of PHM predictions it can be seen that the precision of the PHM predictions was superior to the Weibull baseline.

16 PHM Matlab Toolbox

A number of technical limitations exist during the workflow of designing and testing PHM algorithms. For instance laboratory tests at a single stress level may not be representative of field usage conditions. Many different failure modes can occur in a product than those explored through accelerated life testing. And lastly it is difficult to validate that a PHM algorithm will meet its specified requirements during actual field use. In an effort to increase the ease of building, verifying and validating PHM algorithms a PHM matlab toolbox has been created.

16.1 Introduction

The PHM toolbox is a collection of functions that can be combined with sensors to create PHM implementations. The recursive filtering based algorithms presented in this document were implemented in the toolbox. Many of the illustrative examples found in this document are used to validate the functions in the toolbox. Two levels of use are imagined for the toolbox. In the first scenario the toolbox would be used as a finished product as part of a turn-key solution. In the second scenario the toolbox would form a stable foundation for new and innovative improvements to existing PHM algorithms.

16.1.1 Turn-key Solutions

If a user needs a verified and validated solution that is immediately ready to implement, the PHM toolbox has a variety of complete PHM algorithm implementations. Because the code has been used to develop the work in this document it is easy to estimate the accuracy of the implementation when applied to other applications. The user would implement the code as is into a product knowing that the code has a long history of use and development. The turnkey

approach is readily applicable to feasibility studies on new and existing products and to help decrease the time and cost of accelerated life testing.

16.1.2 Building Block Solutions

If a user needs a strong foundation on which to build new PHM implementations, the PHM toolbox is a verified and validated foundation to begin with. Instead of starting with a blank slate, the PHM toolbox provides code for commonly used routines such as the PHM metrics and components for recursive filtering. The nature of the code is very modular and encourages innovation. In PHM algorithm development, sensor and code development is not easily separated. The PHM toolbox provides quick access for developing baseline measures of success and then can be extended to include any additional functionality that is needed to fuse new sensor developments with the constantly changing software requirements.

16.2 Verification and Validation Scripts

A number of script files exist to verify and validate the correct operation of the toolbox. These script files are heavily commented and can also be considered tutorials for how to use the code in the toolbox to build PHM algorithms. Utilizing code blocks that have been verified in numerous different applications allows the user of the toolbox to focus on more pressing design/development matters in the PHM algorithm workflow.

I. Kalman Filter V&V Scripts

   A. Tracking

      1. Re-entry of a ballistic object

   B. Prognostics

      1. Re-entry of a ballistic object

2. Failure in a BGA (ECTC 2010 data set)

II. Extended Kalman Filter V&V scripts

A. Tracking

1. Re-entry of a ballistic object
2. Tracking frequency of a sine wave
3. Tracking amplitude of a sine wave
4. Tracking a harmonically damped system (simple)
5. Tracking a harmonically damped system (improved)

B. Prognostics

1. Re-entry of a ballistic object
2. Failure in a BGA (ECTC 2010 data set)

III. Particle Filter V&V scripts

A. Tracking

1. Re-entry of a ballistic object
2. Quadratic system simulation

B. Prognostics

1. Re-entry of a ballistic object
2. Failure in a BGA (ECTC 2010 data set)

16.3 Code documentation

For the sake of usability and scalability the PHM toolbox is very well documented. In addition to defining inputs and outputs for each function, the code documentation cross references the code to quickly direct users to relevant examples and other related functions. Matlab contains internal toolset that allow custom toolboxes, like the PHM toolbox, to integrate

documentation in the same manner as built in documentation for functions natively packaged with matlab. The result is quick and efficient access to documentation from the Maltab command window.

17 Conclusion

This document focused on predicting (prognosticating) failure in mission and safety critical electronics. Emphasis was placed on developing a measurement technique known as resistance spectroscopy to monitor the health of a component, and implementing algorithms capable of providing advanced warning of failure that could be used to take mitigating action and avoid catastrophic unplanned failures. A risk based decision making framework provided a statistically defendable method for using failure predictions from prognostic algorithms. A variety of BGA package architectures and environmental conditions were used to stress components to failure. The technique was also shown to work for other electrical components, specifically pin and socket connectors. Simulations known as a prognostic hindecaste were used to estimate the future success of proposed prognostic algorithms. The most important components of the presented work will be reviewed, and a discussion of future work will be discussed.

17.1 Salient Features of the Presented Work

A number of novel contributions are found in this work. As a whole, the contributions represent a paradigm shift in the way that mission and safety critical electronics should be analyzed, designed, tested, and maintained. An in depth understanding of traditional reliability methods, and a practical view of how electronics are actually used in fielded products, motivates the development of improved reliability tools such as prognostic health management. The foundation for the experimental topics, the resistance spectroscopy technique, was used in a very different application than it was originally developed for and has been tailored for implementation in embedded applications. The adoption of risk based decision making, recursive

filtering, and prognostic metrics provided a complete framework that could result in a viable product to meet the actual needs of end users. For example, John Deere Electrical Systems has started to implement the techniques described in this document into developmental products and may one day become one of the first fielded prognostic systems that explicitly monitor electrical components. The PHM toolbox highlights contributions that can be bifurcated from any specific package architecture or test conditions. Many of the insights developed in the creation of the toolbox represent fundamental insights into methods for combining existing knowledge of the behavior of electronic assemblies into a new paradigm for reliability of mission and safety critical electronics. The PHM toolbox also provides insight into areas that are not well developed and will require additional work in the future.

17.1.1 Resistance Spectroscopy for Prognostics

The resistance spectroscopy method was developed by prior authors. The novelty of its use in this work was its application to the problem of condition monitoring for electronics. The application of the technique for a prognostic system was successful for three key reasons. The measurement resolution is superior to many other techniques. The circuit is easy to temperature compensate, and the circuit uses a small footprint on a circuit board. Resistance spectroscopy can be implemented with low frequency circuit boards and therefore is cheaper than many other online monitoring techniques. Typically two of the least reliable components on a circuit board, chip resistors and BGA's are good candidates for monitoring using the RS technique.

Proving the repeatability of the RS technique has been another contribution of this work. The technique has been demonstrated on both SnPb and lead-free alloys. Package architectures tested include controlled collapsed solder balls (high lead, SnPb, SAC305, plastic core), copper column arrays, and micro coil spring arrays. Test environments where the RS technique was

successfully applied include drop/shock, vibration, and simultaneous temperature and vibration testing. RS has been shown to successfully monitor both pristine and thermally aged components. The RS technique was also implemented for monitoring changes in contact resistance for pin and socket electrical connectors. There have been no situations where the RS technique was applied and failed to provide a leading indicator of failure.

17.1.2 User-Centric Prognostic Health Management Framework

Individually risk based decision making, recursive filtering, and prognostic performance metrics were developed by prior authors. Their collective use for mission and safety critical electronics was novel. The combined used of the three techniques allowed prognostic systems that are capable of meeting requirements for products outside of the academic laboratory. In this document risk based decision making guided the development of many of the techniques to ensure that the resulting measurement or algorithm would efficiently meet the needs of end users and not simply appeases the imagination of a researcher. Recursive filtering provides a realistic method for real time tracking and dovetails nicely with physics and model based engineering methods. Lastly the prognostic algorithms provide methods for quantitatively defining the performance of an algorithm, and therefore improving performance. The systematic application of leave one out cross validation provided some of the first published insight into the performance that could be expected from an implemented prognostic algorithm on unseen data sets. In much of the prior literature on prognostics, and specifically prognostics for electronics, the logical fallacy of future information leak undermined the ability to fully quantify the expected future performance of a prognostic technique. The strongest argument for the success of the combined framework is the implementation of the presented techniques by an industrial partner outside of the scope of a funded academic research project.

17.1.3 Application Independent Contributions

The contributions of the PHM Matlab toolbox provides a distinction in the work between contributions that are application specific, such as resistance spectroscopy, and fundamental insights into changes necessary for transitioning from traditional reliability methods to prognostic health monitoring. Recursive filtering has a long history when being applied to tracking and navigation problems. Admittedly some of the initial difficulties when learning about recursive filtering were the highly unfortunate choice of notation used in the recursive tracking literature that originated with Kalman's original optimal tracking paper. The PHM toolbox helps bifurcate the nuances of filtering from the formidable challenges in understanding the physics and mechanics of a system that is degrading. It is envisioned that the overhead involved in developing and testing the recursive algorithms, can be drastically reduced by using the PHM toolbox which will allow the researched to concentrate on the challenging task of developing appropriate models to describe/predict damage propagation.

The PHM algorithms tend to be named after the tracking algorithm that is used to implement them, but the prediction phase of the algorithm has just as much effect on performance as the tracking. An understanding of the specific requirements for propagating damage, and uncertainty into the future are application independent contributions from this work. The joint online tracking and parameter estimation of the Extended Kalman filter typically achieves a superior performance over other baseline tracking implementations, while the particle filter typically has the best baseline architecture for propagating uncertainty into the future. An unwieldy number of variations exist for tracking and prognostic algorithms which could alter the conclusion about the best tracking algorithms for prognostics. The PHM toolbox helps provide

baseline implementations as a foundation for comparing and developing new and advanced algorithms.

17.1.4 Summary of PHM Algorithm Performance

Many factors influence the performance of a PHM implementation including the test type, orientation, and stress levels. PHM metrics have been used to quantify the performance of over 30 tests in this document. Each situation is different, but it can be conservatively estimated that the presented algorithms will converge to within an accuracy of $\pm 20\%$ of the actual RUL after about 50% of the time between the feature vector crossing the detection threshold and the failure of the system. Notable exceptions include the results presented for the particle filter and micro-coil spring interconnects where performance was significantly better than average with convergence after only 25% of the time between detection and failure. The two demonstrations of leave one out cross validation provide useful insight into the expected future performance of the demonstrated methods using a rigorous independent evaluation of performance. In the two separate data sets used for leave one out cross validation, the PHM algorithm converged between 25 and 50% of the time after the detection threshold. Furthermore the average cost function using an equal weighting of precision and accuracy for all the experimental results presented in this document have an average score of 0.71, where a score of zero is optimal. For comparison, the best cost function score obtained from an idealistic synthetic data set is calculated as 0.44.

17.2 Future Work

Prognostics in general, and specifically prognostics for electronics is a very young field and much future work is needed. The state of the art, including many of the aspects introduced in this document, still only provide a partial solution to the challenges inherent in designing and operating mission and safety critical electrical systems. Model (physics) based prognostics, as

opposed to purely data driven (data mining) prognostics, seems like the most reasonable approach moving forward, but challenges still exist.

17.2.1 Initialization Concerns

In model based prognostics initialization of recursive filters can be important, and few methods other than experience currently exist to inform the initialization state. In many systems such as turbines, nuclear reactors, and implanted biomedical devices, run to failure data is not available because the cost of catastrophic failure of a system, even in a controlled environment, is too expensive to bear. The limitations of experience for many systems makes initializing prognostic algorithms difficult.

17.2.2 Assumptions about Future Usage

During the damage propagation and failure prediction phase assumptions about future usage conditions are important and difficult to defend. In the special case of accelerated life testing, future usage conditions are tightly controlled and predictable. There is a strong argument that the state of the art prognostic algorithms could be used to reduce the time and cost of accelerated life testing programs in product development activities. In fielded applications future usage conditions are usually less reliably controlled and many challenges exist for quantifying observations and making assumptions about future usage.

17.2.3 Defendable Uncertainty Quantification

The challenges associated with assumptions about future usage of a system leads to a discussion of the challenges associated with quantifying uncertainty in a defendable manner. With such limited information on fielded prognostic implementations, it is hard to quantify if the current methods for quantifying uncertainty are acceptable. In mission and safety critical systems

low probability events are very important to understand and mitigate. Errors in uncertainty quantification using traditional reliability methods are known to result in unbearable failures. It could be reasoned that uncertainty quantification in prognostics will undermine the effectiveness of PHM unless improvements can be realized in the future. As discussed earlier, the best combination of tracking and damage propagation has not even been clearly identified, with EKF and PF based implementation both having some merit.

17.2.4 Challenges in the Design, Verification & Validation of PHM Implementations

Another inherent challenge for prognostics is the design, verification and validation of prognostic implementations. Accelerated life testing is popular for product design because it is timely and cost effective compared to actual usage tests which could last for years or decades. State of the art PHM techniques could be applied directly to predicting failures in ALT's. Unfortunately the needed to validate the correct operation of a fielded PHM implementation does not yet have a clear analogy to ALT because future loading conditions are not explicitly known in advance. In many cases the cost of run to failure testing is unbearable even in a controlled environment and therefore leaves a designer with little justifiable motivation for choosing one design over another. Even if a system is not considered mission or safety critical, the intangible damage caused by experimenting with the reliability of a paying customers products is difficult to justify. Currently only in a few select cases is there enough information and experience to develop PHM implementations for a select few components.

18 References

Al-Yahyai, S., Charabi, Y., Gastli, A., Review of the Use of Numerical Weather Prediction (NWP) Models for Wind Energy Assessment, Renewable and Sustainable Energy, Vol. 14, No. 9, pp. 3192–3198, 2010.

Alpaydin, E. Introduction to machine learning. The MIT Press, 2004.

American Society of Mechanical Engineers (ASME), Guide for Verification and Validation in Computational Solid Mechanics, American Society of Mechanical Engineers, 2006.

Angadi, S. V., Wilson, W. E., Jackson, R. L., Flowers, G. T. and Rickett, B. I., A Multi-Physics Finite Element Model of an Electrical Connector Considering Rough Surface Contact., IEEE Holm Conference on Electrical Contacts, 2008.

Arulampalam, M.S., S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on Signal Processing, 50, no. 2, 174-188, 2002.

Baldwin, C., J. Kiddy, T. Salter, P. Chen, and J. Niemczuk, Fiber Optic Structural Health Monitoring System: Rough Sea Trials Testing of the RV Triton, MTS/IEEE Oceans 2002, Volume 3, pp. 1807-1814, 2002.

Balchen, J., Jenssen, N., Mathisen, E., A Dynamic Positioning system based on Kalman Filtering and Optimal Control, Modeling, Identification and Control, Vol 1, No. 3, pp.135–163, 1980.

Banks, J., Merenich, J., Cost Benefit Analysis for Asset Health Management Technology, IEEE Annual Reliability and Maintainability Symposium (RAMS), 2007.

Banks, J., Reichard, K., Crow, E., Nickell, K., How Engineers Can Conduct Cost-Benefit Analysis for PHM Systems, IEEE Aerospace and Electronic Systems Magazine, Vol. 24, no. 3, pp. 22-30, 2009.

Banyasz, C., Keviczky, L,. Identification and System Parameter Estimation, International Federation of Automatic Control, and International Federation of Operational Research Societies, 1992.

Bar-Shalom, Yaakov, X. Rong Li, Xiao-Rong Li, and Thiagalingam K., Estimation with Applications to Tracking and Navigation. John Wiley and Sons, 2001.

Barke, D., Chiu, W., K., Structural Health Monitoring in the Railway Industry: A Review, Structural Health Monitoring, Vol. 4, No. 1, pp. 81-93, 2005.

Batra, A., Fang, L., Constable, J., Implementation of Low-cost Failure Detection System Using Resistance Spectroscopy, Electronic Components and Technology Conference, 2003.

Batzel, T., Swanson, Prognostic Health Management of Aircraft Power Generators, IEEE Transactions on Aerospace and Electronic Systems, Vol. 45, No. 2, pp. 473-482, 2009.

Bechhoefer, E., Kingsley, M., Menon, P., Bearing Envelope Analysis Window Selection Using Spectral Kurtosis Techniques, IEEE International Conference on Prognostics Health Management, 2011.

Belle, G., Fisher, L., Heagerty, P., Lumley, T., Biostatistics: A methodology for Health Sciences, Wiley, 2nd Edition, 2004.

Bernieri, A., D'Apuzzo, M., Sansone, L., Savastano, M., A Neural Network Approach for Identification and Fault Diagnosis on Dynamic Systems, IEEE Transactions on Instrumentation and Measurement, Vol. 43, No. 6 pp. 867–873, 1994.

Bevly, D. M, and B. Parkinson. Cascaded Kalman filters for accurate estimation of multiple biases, dead-reckoning navigation, and full state feedback control of ground vehicles. IEEE Transactions on Control Systems Technology, Vol. 15, No. 2, pp. 199–208, 2007.

Bhate, D., Chan, D., Subbarayan, G., Chiu, T., Gupta, V., Edwards, D., Constitutive Behavior of Sn3.8Ag0.7Cu and Sn1.0Ag0.5Cu Alloys at Creep and Low Strain Rate Regimes IEEE Transactions on Components and Packaging Technologies, Vol. 31, No. 3, pp 622–633, 2008.

Black, F., Scholes, M., The Pricing of Options and Corporate Liabilities, Journal of Political Economy, 81, no. 3, pp. 637-654, 1973.

Blackwell, G.R. The Electronic Packaging Handbook. CRC, 2000.

Bréhard, T., Cadre, J., Initialization of Particle Filter and Posterior Cramér-Rao Bound for Bearings-only Tracking in Modified Polar Coordinate System, Technical Report of the Institute for Information Research and Automation, 2004.

Bromba, Manfred, Ziegler, H., Application Hints for Savitzky-Golay Digital Smoothing Filters, Analytical. Chemistry, Vol.. 53, No. 11, pp. 1583–1586, 1981.

Brown, D., M. Abbas, A. Ginart, I. Ali, P. Kalgren, and G. Vachtsevanos. Turn-Off Time as an Early Indicator of Insulated Gate Bipolar Transistor Latch-up, IEEE Transactions on Power Electronics, 2011.

Bryant, M. D.,Resistance Buildup in Electrical Connectors Due to Fretting Corrosion of Rough Surfaces, IEEE Transactions On Components, Packaging and Manufacturing Technology – Part A, Vol. 17, No. 1 pp. 86-95, 1994.

Butler, W., Constable, J., Pitarresi, J., Resistance Spectroscopy Applied to Accelerated Life Testing of Solder Joints, Confernce of the Society of Experimental Mechanics, 2000.

Byer, B., Hess, A., Fila, L., Writing a convincing cost benefit analysis to substantiate autonomic logistics, IEEE Aerospace Conference, 2001.

Cadini, F., Zio, E., Avram, D., Monte Carlo-based filtering for fatigue crack growth estimation, Probabilistic Engineering Mechanics, Vol. 24, No. 3, pp. 367-373, 2009.

Cappé, O., S. J Godsill, and E. Moulines. An overview of existing methods and recent advances in sequential Monte Carlo. Proceedings of the IEEE, Vol. 95, No. 5, pp. 899–924, 2007.

Carrasco, M., Cassady, C., A Study of the Impact of Prognostic Errors on System Performance, IEEE Annual Reliability and Maintainability Symposium (RAMS) , 2006.

Cartwright, R., Paynter, R.,Hills, D., Nowell, D., Long Term Wear of Complete Contacts Subject to Fretting, Wear, Vol. 271, No. 11–12, pp. 2821–2825, 2011.

Chang, P., C., Flatau, A., and Liu, S.,C., Review Paper: Health Monitoring of Civil Infrastructure, Structural Health Monitoring, Vol.2, No.3, pp. 257-267, 2003.

Chong, D., Che, F., Pang, J., Ng, K., Jane, Y., Tan, P., Drop impact reliability testing for lead-free and lead-based soldered IC packages, Microelectronics and Reliability, Vol. 46, No. 7, pp.1160-1171, July 2006.

Clech, J., Manock, J., Noctor, D., Bader, F., Augis, J., A Comprehensive Surface Mount Reliability Model Covering Several Generations of Packaging and Assembly Technology, IEEE Transactions On Components, Hybrids, and Manufacturing Technology, Vol. 16, No. 8, pp 949–960, 1993.

Clerc, M., Kennedy, J., The Particle Swarm - Explosion, Stability, and Convergence in a Multidimensional Complex Space, IEEE Transactions on Evolutionary Computation, 2002.

Coble, J., Humberstone, M., Hines,W., Adaptive Monitoring, Fault Detection and Diagnostics, and Prognostics System for the IRIS Nuclear Plant, Annual Conference of the Prognostics and Health Management Society ,2010.

Constable, J., Lizzul, C., An Investigation of Solder Joint Fatigue Using Electrical Resistance Spectroscopy, IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part A, Vol. 18, No. 1 pp. 142–152, 1995.

Constable, J., Use of Interconnect Resistance as a Reliability Tool. Electronic Components and Technology Conference, 1994.

Daigle, M., Goebel, K., Model-based Prognostics with Fixed-lag Particle Filters. Conference of the PHM Society, 2009.

Daigle, M., Goebel, K., Model-based prognostics under limited sensing. IEEE Aerospace Conference, 2010.

Daniel, C. and Mucklich, F., Electrical Behavior of periodically microstructered SN/CuSn4 Contact Models under Fretting Conditions, Wear, Vol. 257, pp. 266-270, 2004.

Dally, J., Lall, P., Suhling, J., Mechanical Design of Electronic Systems. College House Enterprises, LLC, 2008.

Dally, James, Experimental Stress Analysis, College House Enterprises, 2005.

Darveaux, R., Effect of Simulation Methodology on Solder Joint Crack Growth Correlation and Fatigue Life Prediction, Journal of Electronic Packaging, Vol. 124, No. 3 pp. 147–154, 2002.

Darveaux, R., C. Reichman, and N. Islam, Interface failure in lead free solder joints, 56th Electronic Components and Technology Conference, San Diego, CA, pp. 906-917, 2006.

Datar, V., Matthews, S., European Real Options: An Intuitive Algorithm for the Black-Scholes Formula. Journal of Applied Finance, Vol. 14, pp 45–51, 2004.

Dawson, M. Practical Lessons in Actuarial Science: An Elementary Text-book, The Spectator Company, 1898. (not a typo, year of publication was 1898)

De Souza, D., Lamberson, L., Bayesian Weibull Reliability Estimation, IIE Transactions, Vol. 27, No. 3, pp 311–320, 1995.

Doucet, A., De Freitas, N., Gordon, N., Sequential Monte Carlo methods in practice. Springer Verlag, 2001.

Downes, Clive G., ASTRAEA T7 — An architectural outline for system health management on civil UAVs, Institution of Engineering and Technology Conference on Autonomous Systems, 1-4, 2007.

Elder IV, J., Top Ten Data Mining Mistakes and How to Avoid Them, Salford Systems Data Mining Conference, 2005.

Elerath, J., Shah, S., Server Class Disk Drives: How Reliable Are They?, Reliability and Maintainability, 2004.

Engel, S., PHM Engineering Perspectives, Challenges and 'Crossing the Valley of Death, Annual Conference of the Prognostics and Health Management Society, San Diego, CA, pp. 1-21, 2009.

Evensen, G., Van Leeuwen, P., Assimilation of Geosat Altimeter Data for the Agulhas Current Using the Ensemble Kalman Filter with a Quasi-geostrophic Model, Monthly Weather, 1994.

Evensen, Geir, The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation, Ocean Dynamics, Vol. 53, No. 4, pp. 343–367, 2003.

Farina, A., Ristic, B., Timmoneri, L., Cramer-Rao Bound for Nonlinear Filtering with Pd<1 and Its Application to Target Tracking, IEEE Transactions on Signal Processing, Vol. 50, No. 8 pp 1916–1924, 2002.

Feldman, K., Jazouli, T., Sandborn, P., A Methodology for Determining the Return on Investment Associated With Prognostics and Health Management, IEEE Transactions on Reliability, Vol. 58, no. 2, pp.305-316, 2009.

Flowers, G. T., Xie, F., Bozack, M. J., and Malucci, R. D., Vibration Thresholds for Fretting Corrosion in Electrical Connectors, IEEE Transactions On Components and Packaging Technology, Vol. 27, No. 1, pp. 583-590, 2004.

Flowers, G. T., Xie, F., Bozack, M. J., Horvath, R., Malucci, R. D. and Rickett, B. I., Modeling Early Stage Fretting of Electrical Connectors Subjected to Random Vibration, IEEE Transactions On Components and Packaging Technology, Vol. 28, No. 4, pp. 721-727, 2005.

Flowers, G. T., Xie, F., Bozack, M. J., Hai, X., Rickett, B. I., Malucci, R. D., A Study of the Physical Characteristics of Vibration-Induced Fretting Corrosion, IEEE Transactions On Components and Packaging Technology, Vol. 29, No. 24, pp. 318-325, 2006.

Fouvry, S., Jedrzejczyk, P., Chalandon, P., Introduction of an Exponential Formulation to Quantify the Electrical Endurance of Micro-contacts Enduring Fretting Wear: Application to Sn, Ag and Au Coatings, Wear, Vol. 271, No. 9–10, pp. 1524–1534, 2011.

Fu, R., Choe, S., Jackson, R., Flowers, G., Bozack, M., Zhong, L., Kim, D., Experimental Study of the Vibration-Induced Fretting of Silver-Plated High Power Automotive Connectors, Holm Conference On Electrical Contacts, 2010.

Fulton, W., WeibullSmith User Guide, Fulton Findings LLC, P.O. Box 6517, San Pedro, CA 90734-6517, 2002

Goebel, K., Saha, B., Saxena, A., Celaya, J., Christophersen, J., Prognostics in Battery Health Management, IEEE Instrumentation & Measurement Magazine, no. 4, 2008, pp. 33-40, 2008.

Goodman, D., Wood, S., Turner, A., Return-on-Investment (ROI) for Electronic Prognostics in Mil/Aero systems, IEEE Autotestcon, 2005.

Gorry, P., General Least-squares Smoothing and Differentiation by the Convolution (Savitzky-Golay) Method, Analytical Chemistry, Vol. 62, No. 6, pp 570–573,1990.

Grewal, M. S, Kain, J., Kalman Filter Implementation With Improved Numerical Properties, IEEE Transactions On Automatic Control, Vol. 55, No. 9, pp. 2058–2068, 2010.

Grubic, T., Jennions, I., Baines, T., The Interaction of PSS and PHM - A mutual benefit case, Annual Conference of the Prognostics and Health Management Society, 2009

Gueler, G., Modelling, Design and Analysis of an Autopilot for Submarine Vehicles, International Shipbuilding Progress, Vol. 36, No. 405, pp. 51–85, 1989.

Gustafsson, F., Bergman, U. Forssell, J. Jansson, R. Karlsson, Nordlund, P., Particle filters for positioning, navigation, and tracking. IEEE Transactions on signal processing 50, no. 2, pp. 425–437, 2002.

Hallinan, A., Jr., A Review of the Weibull Distribution, Journal of Quality Technology, Vol. 25, No. 2, pp. 85-93, 1993.

Hamada, Michael S. Bayesian Reliability. Springer, 2008.

Hao, L., Jinsong, Y., Ping, Z., Xingshan, L., Review on Verification and Validation Technology in Integrated Health Management System, International Conference on Electronic Measurement & Instruments, 2011.

Harper, C., Electronic Packaging and Interconnection Handbook. McGraw-Hill Professional, 2005.

Hayward, R., Gebre-Egziabher, D., Schwall, M., Powell, J., Wilson, J.,. Inertially Aided GPS Based Attitude Heading Reference System (AHRS) for General Aviation Aircraft. Proceedings of the Institute of Navigation ION-GPS Conference, pp. 1415–1424, 1997.

He, David, Manon, P., Li, R., Seçkiner, S., Bechhoefer, E. Gear Fault Location Detection for Split Torque Gearbox Using AE Sensors., Annual Conference of the PHM Society, 2010.

Heredia-Zavoni, E., Santa-Cruz, S., Maintenance decisions for offshore structures using real options theory, ASME International Conference on Offshore Mechanics and Artic Engineering, 2004.

Herring, Kenneth D, and Frank Seiler, Evaluation of an Estimator for Real-Time Missile Tracking. Defense Technical Information Center, 1974.

Hibbeler, R,. Mechanics of Materials. Pearson Education, 2003.

Higham, Desmond., An Introduction to Financial Option Valuation, Cambridge University Press, 2004.

Hills., R., Dowding, K., Statistical Validation of Engineering and Scientific Models: Bounds, Calibration, and Extrapolation by, SANDIA REPORTS, 2005.

Hongtao M., Suhling, J., Lall, P., Bozack, M., Reliability of the Aging Lead Free Solder Joint, IEEE Electronic Components and Technology Conference, 2006.

Hongtao M., Suhling, J., Lall, P., Bozack, M., The Influence of Elevated Temperature Aging on Reliability of Lead Free Solder Joints. IEEE Electronic Components and Technology Conference, 2007.

Hongtao M., Suhling, J., A Review of Mechanical Properties of Lead-free Solders for Electronic Packaging, Journal of Materials Science, Vol. 44, pp.1141–1158, 2009.

Hopp, W., Spearman, M., Factory Physics. Waveland Press Inc., 2011.

Hull, John., Options, Futures, and Other Derivatives, Prentice Hall, 4th Edition, 2000.

Hutapea, P., Grenestedt, J., Effect of Temperature on Elastic Properties of Woven-glass Epoxy Composites for Printed Circuit Board Applications, Journal of Electronic Materials, Vol. 32, no. 4, pp. 221–227, 2003.

IPC, IPC-SM-785: Guidelines for Accelerated Reliability Testing of Surface Mount Solder Attachments, 2215 Sanders Road, Northbrook, IL 60062-6135, 1992.

Isermann, R., Ballé, P., Trends in the Application of Model-based Fault Detection and Diagnosis of Technical Processes, Control Engineering Practice, Vol. 5, No. 5, pp. 709–719, 1997.

Jacklin, S., Lowry, M., Schumann, J., Gupta, P., Bosworth, J., Zavala, E., Kelly, J., Hayhurst, H., Belcastro, C., Verification, Validation, and Certification Challenges for Adaptive Flight-critical Control System Software, Proceedings of American Institute of Aeronautics and Astronautics AIAA Guidance Navigation and Control Conference and Exhibit, American Institute of Aeronautics and Astronautics, 2004.

Jacklin, S., Schumann, J., Gupta, P., Richard, R., Guenther, K., Soares, F., Development of Advanced Verification and Validation Procedures and Tools for the Certification of Learning Systems in Aerospace Applications, Proceedings of Infotech@ Aerospace Conference, 2005.

Jarrell, D., Sisk, D., Bond, L., Prognostics and Condition Based Maintenance (CBM) A Scientific Crystal Ball, International Congress on Advanced Nuclear Power Plants, 2002.

JEDEC Standard JESD22-B111. Board level drop test method of components for handheld electronic products. JEDEC Solid State Technology Assoc; 2003.

JEDEC Solid State Technology Association, JESD22-B103B: Vibration, Variable Frequency, 2500 Wilson Boulevard, Arlington, VA 22201-3834, 2006.

JEDEC Solid State Technology Association, JESD22-B111: Board Level Drop Test Method of Components for Handheld Electronic Products, 2500 Wilson Boulevard, Arlington, VA 22201-3834, 2006.

Jedrzejczyk, P., Fauvry, S. and Chalandon, P.,A Fast Methodology to Quantify Electrical-Contact Behavior Under Fretting Loading Conditions, Wear, Vol. 267, pp. 1731-1740, 2009.

Jiang, M., Dummer, D., Bayesian Reliability Demonstration Test in a Design for Reliability Process, Reliability and Maintainability Symposium, 2009.

Kalman, R., A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME Journal of Basic Engineering, Vol. 82, No. D, pp. 35-45, 1960.

Kapurch, S., NASA Systems Engineering Handbook, DIANE Publishing, 2010.

Kennedy, J., Eberhart, R., Particle Swarm Optimization, IEEE International Conference on Neural Networks, 1995.

Kleinbaum, D., Klien, M., Pryor, E., Logistic Regression: a Self-learning Text. Springer, 2002.

Kleinbaum, D., Klien, M., Survival Analysis: A self learning text, Springer, 2nd Edition, 2005

Klien, J., Goel, P., Survival Analysis: State of the Art, Kluwer Academic Publishers, 1991

Kim, J., Wishart, S., Sukkarieh, S., Real-time Navigation, Guidance and Control of a UAV Using Low-cost Sensors, International Conference of Field and Service Robotics, 2003.

Ko, J., Ni., Yi., Technology Developments in Structural Health Monitoring of Large-Scale Bridges. Engineering Structures , no. 12, pp. 1715-1725, 2005.

Kreyszig, E., Advanced Engineering Mathematics, John Wiley and Sons, 2010.

Krohling, R. A., Gaussian Swarm: a Novel Particle Swarm Optimization Algorithm, IEEE Conference on Cybernetics and Intelligent Systems, 2004.

Kuciauskas, A., NexSat: Reaching out to users within and outside of the US, Conference on Satellite Meteorology and Oceanography, 2010.

Kulkarni, C., Biswas, G., Koutsoukos, X., Celaya, J., Goebel, K., Experimental Studies of Ageing in Electrolytic Capacitors, Annual Conference of the Prognostics and Health Management Society, 2010.

Lall, P., Estimating Influence of Temperature on Microelectronic Device Reliability: Physics of Failure Approach, CRC Press, 1996.

Lall, P., Islam, N., Rahim, K., Suhling, J., Gale, S., Leading Indicators-of-Failure for Prognosis of Electronic and MEMS Packaging, 54th Electronics Components and Technology Conference, Las Vegas, Nevada, June 1 – 4, 2004a.

Lall, P., Islam, N., Shete, T., Evans, J., Suhling, J., Gale, S., Damage Mechanics of Electronics on Metal-Backed Substrates in Harsh Environments, 54th Electronic Components and Technology Conference, Las Vegas, Nevada, June 1 - 4, 2004b.

Lall, P., N. Islam, J. C. Suhling, and R. Darveaux, Model for BGA and CSP Reliability in Automotive Underhood Applications, IEEE Transactions on Components and Packaging Technologies, Volume 27, Number 3, pp. 585-593, 2004c.

Lall, P., D. Panchagade, Y. Liu, R. W. Johnson, and J. C. Suhling, Models for Reliability Prediction of Fine-Pitch BGAs and CSPs in Shock and Drop-Impact, Proceedings of the 54th Electronic Components and Technology Conference, pp. 1296-1303, Las Vegas, NV, June 1-4, 2004d.

Lall, P., Islam, N., Choudhary, P., Suhling, J., Prognostication and Health Monitoring of Leaded and Lead Free Electronic and MEMS Packages in Harsh Environments, Proceedings of the 55th IEEE Electronic Components and Technology Conference, pp. 1-9, Orlando, FL, June 1-3, 2005a.

Lall, P., Panchagade, D., Choudhary, P., Suhling, J., Gupte, S., Failure-Envelope Approach to Modeling Shock and Vibration Survivability of Electronic and MEMS Packaging, Proceedings of the 55th IEEE Electronic Components and Technology Conference, Orlando, FL, pp. 480 – 490, June 1 – 3, 2005b.

Lall, P., Gupte, S., Choudhary, P., Suhling, J., Solder-Joint Reliability in Electronics Under Shock and Vibration using Explicit Finite Element Sub-modeling, Proceedings of the 56th IEEE Electronic Components and Technology Conference, San Diego, California, pp.428-435, May 30-June 2, 2006a.

Lall, P., Choudhary, P., Gupte, S., Suhling, J., Health Monitoring for Damage Initiation & Progression during Mechanical Shock in Electronic Assemblies, Proceedings of the 56th IEEE Electronic Components and Technology Conference, San Diego, California, pp.85-94, May 30-June 2, 2006b.

Lall, P., Hande, M., Singh, N., Suhling, J., Lee, J., Feature Extraction and Damage Data for Prognostication of Leaded and Leadfree Electronics, Proceedings of the 56th IEEE Electronic Components and Technology Conference, San Diego, California, pp.718-727, May 30-June 2, 2006c.

Lall, P., Islam, N., Rahim, K., Suhling, J., Gale, S., Prognostics and Health Management of Electronic Packaging, IEEE Transactions on Components and Packaging Technologies, Volume 29, Number 3, pp. 666-677, September 2006d.

Lall, P., Islam, N., Shete, T., Evans, J., Suhling, J., Gale, S., Damage Mechanics of Electronics on Metal-Backed Substrates in Harsh Environments, IEEE Transactions on Components and Packaging Technologies, Volume 29, Number 1, pp. 204-212, March 2006e.

Lall, P., D. Panchagade, Y. Liu, R. W. Johnson, and J. C. Suhling, Models for Reliability Prediction of Fine-Pitch BGAs and CSPs in Shock and Drop-Impact, IEEE Transactions on Components and Packaging Technologies, Volume 29, Number 3, pp. 464-474, September 2006f.

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 Electronics Components and Technology Conference, Reno, Nevada, pp. 1161-1178, May 30-June 1, 2007a.

Lall, P., Gupte, S., Choudhary, P., Suhling, J., Solder-Joint Reliability in Electronics Under Shock and Vibration using Explicit Finite Element Sub-modeling, IEEE Transactions on Electronic Packaging Manufacturing, Volume 30, No. 1, pp. 74-83, January 2007b.

Lall, P., M. Hande, C. Bhat, J. Suhling, Jay Lee, Prognostics Health Monitoring (PHM) for Prior-Damage Assessment in Electronics Equipment under Thermo-Mechanical Loads, Electronic Components and Technology Conference, Reno, Nevada, pp. 1097-1111, May 29 – June 1, 2007c.

Lall, P. Panchagade, D., Liu, Y., Johnson, W., Suhling, J., Smeared Property Models for Shock-Impact Reliability of Area-Array Packages, ASME Journal of Electronic Packaging, Volume 129, pp. 373-381, December 2007d.

Lall, P., Hande, M., Bhat, C., Islam, N., Suhling, J., Lee, J., Feature Extraction and Damage-Precursors for Prognostication of Lead-Free Electronics, Microelectronics Reliability, Volume 47, pp. 1907–1920, December 2007e.

Lall, P., Gupte, S., Choudhary, P., Suhling, J., Solder Joint Reliability in Electronics Under Shock and Vibration Using Explicit Finite-Element Submodeling, IEEE Transactions on Electronics Packaging Manufacturing, Vol. 30, No. 1, pp. 74–83, 2007f.

Lall, P., Choudhary, P., Gupte, S., Suhling, J., Health Monitoring for Damage Initiation and Progression during Mechanical Shock in Electronic Assemblies, IEEE Transactions on Components and Packaging Technologies, Vol. 31, No. 1, pp. 173-183, March 2008a.

Lall, P., Panchagade, D., Choudhary, P., Gupte, S., Suhling, J., Failure-Envelope Approach to Modeling Shock and Vibration Survivability of Electronic and MEMS Packaging, IEEE Transactions on Components and Packaging Technologies, Vol. 31, No. 1, pp. 104-113, March 2008b.

Lall, P., Hande, M., Bhat, C., More, V., Vaidya, R., Suhling, J., Algorithms for Prognostication of Prior Damage and Residual Life in Lead-Free Electronics Subjected to Thermo-Mechanical Loads, Proceedings of the 10th Intersociety Thermal and Thermo-mechanical Phenomena (ITherm), Orlando, Florida, pp. 638-651, May 28-31, 2008c.

Lall, P., Bhat, C., Hande, M., More, V., Vaidya, R, Pandher, R., Suhling, J., Goebel, K., Interrogation of System State for Damage Assessment in Lead-free Electronics Subjected to Thermo-Mechanical Loads, Proceedings of the 58th Electronic Components and Technology Conference (ECTC), Orlando, Florida, pp. 918-929, May 27-30, 2008d.

Lall, P., Iyengar, D., Shantaram, S., S., Gupta, P., Panchagade, D., Suhling, J., KEYNOTE PRESENTATION: Feature Extraction and Health Monitoring using Image Correlation for Survivability of Leadfree Packaging under Shock and Vibration, Proceedings of the 9th International Conference on Thermal, Mechanical, and Multi-Physics Simulation and Experiments in Micro-Electronics and Micro-Systems (EuroSIME), Freiburg, Germany, pp. 594-608, April 16-18, 2008e.

Lall, P., Iyengar, D., Shantaram, S., Pandher, R., Panchagade, D., Suhling, J., Design Envelopes and Optical Feature Extraction Techniques for Survivability of SnAg Leadfree Packaging Architectures under Shock and Vibration, Proceedings of the 58th Electronic Components and Technology Conference (ECTC), Orlando, Florida, pp. 1036-1047, May 27-30, 2008f.

Lall, P., R. Lowe, and K. Goebel, Resistance spectroscopy-based condition monitoring for prognostication of high reliability electronics under shock-impact, 59th Electronic Components and Technology Conference, San Diego, CA, pp. 1245-1255, 2009a.

Lall, P., Lowe, R., Suhling, J., Prognostication Based On Resistance-Spectroscopy For High Reliability Electronics Under Shock-Impact, ASME IMECE, pp. 1-12, Lake Buena Vista, FL, Nov 13-19, 2009b.

Lall, P., Lowe, R., Goebel, K. Suhling, J., Leading-Indicators Based on Impedance Spectroscopy for Prognostication of Electronics under Shock and Vibration Loads, In International Conference and Exhibition on Packaging and Integration of Electronic and Photonic Systems, MEMS, and NEMS (ASME InterPack'09), San Francisco, CA, 2009c.

Lall, P., Lowe, R., Goebel, K. Suhling, J., Prognostication for Impending Failure in Leadfree Electronics Subjected to Shock and Vibration Using Resistance Spectroscopy, International Symposium on Microelectronics, IMAPS, 2009d.

Lall, P., Gupte,S., Choudhary, P., Darveaux, R., Cohesive-Zone Explicit Submodeling for Shock Life-Prediction in Electronics, IEEE Transactions On Components and Packaging Technologies, Vol 32, No. 2, pp. 365–377, 2009e.

Lall, P., R. Lowe, and K. Goebel. Prognostics Using Kalman-Filter Models and Metrics for Risk Assessment in BGAs Under Shock and Vibration Loads. 60th Electronic Components and Technology Conference 2010a.

Lall, P., R. Lowe, and K. Goebel. Use of Prognostics in Risk-Based Decision Making for BGAs Under Shock and Vibration Loads. Thermal and Thermomechanical Phenomena in Electronics Systems, 2010b.

Lall, Pradeep, Mandar Kulkarni, Arjun Angral, Dhananjay Panchagade, and Jeff Suhling. Digital-image correlation and XFEM based shock-reliability models for leadfree and advanced interconnects. Proceedings of the 60th Electronic Components and Technology Conference, 91-105, 2010c.

Lall, P., Gupta, P., Kulkarni, M., Hofmeister, J., Time–Frequency and Autoregressive Techniques for Prognostication of Shock-Impact Reliability of Implantable Biological Electronic Systems, IEEE Transactions On Electronics Packaging Manufacturing, Vol. 33, No. 4 pp. 289–302, 2010d.

Lall, P., Gupta, P., Panchagade, D., Self-organized Mapping of Failure Modes in Portable Electronics Subjected to Drop and Shock, Electronic Components and Technology Conference, 2010e.

Lall. P., More, V., Vaidya, R., Goebel, K., Assessment of Residual Damage in Lead-free Electronics Subjected to Multiple Thermal Environments of Thermal Aging and Thermal Cycling, Proceedings of 60th ECTC, pp. 206-218, Las Vegas, Nevada, USA, June 2-5, 2010f.

Lall, P., Kulkarni, M., Angral, A., Panchagade, D., Suhling, J., Digital-image Correlation and XFEM Based Shock-reliability Models for Leadfree and Advanced Interconnects, Electronic Components and Technology Conference (ECTC), pp. 91–105, 2010g.

Lall, P., Shantaram, S., Panchagade, D., Peridynamic-models Using Finite Elements for Shock and Vibration Reliability of Leadfree Electronics, IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2010h.

Lall, P., Lowe, R., Goebel, K., Extended Kalman Filter Models and Resistance Spectroscopy for Prognostication and Health Monitoring of Leadfree Electronics Under Vibration, IEEE International Conference on Prognostics and Health Management, 2011a.

Lall, P., Lowe, R., Goebel, K., Particle Filter Models and Phase Sensitive Detection for Prognostication and Health Monitoring of Leadfree Electronics Under Shock and Vibration, IEEE Electronic Components and Technology Conference, 2011b.

Lall, P., Shantaram, S., Kulkarni, M., Limaye, G., Suhling, J., High Strain-rate Mechanical Properties of SnAgCu Leadfree Alloys, Electronic Components and Technology Conference (ECTC), 2011c.

Lam, L., Maul, C. and McBride, J. W., Temperature, Humidity and Pressure Measurement on Automotive Connectors, IEEE Transactions On Components and Packaging Technology, Vol. 29, No. 2, pp. 333-340, 2006.

Lee, W., Nguyen, L., Selvaduray, G., Solder Joint Fatigue Models: Review and Applicability to Chip Scale Packages, Microelectronics Reliability, Vol. 40, No. 2, 2000.

Lizzul, C., Constable, J.,Westby, G., Fatigue Investigation of Lap Shear Solder Joints Using Resistance Spectroscopy, Electronic Components and Technology Conference, 1994.

Lorenz, E., A Study of the Predictability of a 28-variable Atmospheric Model, Tellus, Vol. 17, No. 3 pp. 321–333, 1965.

Luna, J., Kolodziejski, P., Frankle, N., Conroy, D., and Shroder, R., Strategies For Optimizing The Application Of Prognostic Health Management To Complex Systems, Machine Failure Prevention Technology Conference , 2009.

Ma, H., Suhling, J., Zhang, Y., Lall, P., Bozack, M., The Influence of Elevated Temperature Aging on Reliability of Lead Free Solder Joints, Electronic Components and Technology Conference (ECTC) , 2007.

Ma, H., Suhling, J., A Review of Mechanical Properties of Lead-free Solders for Electronic Packaging. Journal of Materials Science, Vol. 44, January, pp. 1141-1158, 2009.

Marur, Prabhakar R., and Hareesh V. Tippur. A strain gage method for determination of fracture parameters in bimaterial systems. Engineering Fracture Mechanics 64, no. 1, 87-104 1999.

Mathews, S., Datar, V., Johnson, B., A Practical Method for Valuing Real Options: The Boeing Approach, Journal of Applied Corporate Finance, Vol. 19, no. 2, pp. 95-104, 2007.

Mathworks, Statistics Toolbox: User's Guide (r2011b). Retrieved February 15th, 2012 from http://www.mathworks.com/help/toolbox/stats/wblfit.html

Matkowski, P., Zawierta, R., Felba, J., Characterization of PCB Materials Vibration Response of Printed Circuit Board in Wide Range of Temperature. IEEE International Spring Seminar on Electronics Technology(ISSE), 2009.

Mattila, T., Laurila, L., Kivilahti, J., Metallurgical Factors Behind the Reliability of High-Density Lead-Free Interconnections, Micro-and Opto-Electronic Materials and Structures: Physics, Mechanics, Design, Reliability, Packaging, pp. 313–350, 2005.

Mattila, T., Kivilahti, J., Reliability of Lead-free Interconnections Under Consecutive Thermal and Mechanical Loadings, Journal of Electronic Materials, Vol. 35, no. 2, pp. 250-256, 2006.

Mattila, T., Li, J., Kivilahti, J., On the Effects of Temperature on the Drop Reliability of Electronic Component Boards, Microelectronics Reliability, In Press, 2011.

Minges, M., Electronic Materials Handbook: Packaging. Vol. 1. CRC, 1989.

Mishra, S., Pecht, M., Smith, T., McNee, I., Harris, R., Remaining Life Prediction of Electronic Products Using Life Consumption Monitoring Approach, Proceedings of the European Microelectronics Packaging and Interconnection Symposium, 2002.

Mishra, S., Ganesan, S., Pecht, M., Xie, J., Life Consumption Monitoring for Electronics Prognostics, IEEE Aerospace Conference, 2004.

Mohanty, J.R., B.B. Verma, and P.K. Ray. Prediction of fatigue crack growth and residual life using an exponential model: Part I (constant amplitude loading). International Journal of Fatigue, 31, no. 3, March: 418-424, 2009.

Moore, G., Cramming More Components Onto Integrated Circuits, Electronics Magazine, Vol. 38, No. 8, 1965.

Moriguchi, S, Hayashi, Y., Nose, Y., Maehara, Y., Korenaga, D., Sugimachi, K.," Comparison of the Logistic Regression and the Cox Proportional Hazard Models in Retrospective Studies on the Prognosis of Patients with Gastric Cancer, Journal of Surgical Oncology, Vol. 52, No. 1, 1993.

Munns, T., Kent, R., Structural Health Monitoring: Degradation Mechanism and System Requirements, Digital Avionics Systems Conferences,2000.

Mustafa, M., Cai, Z., Suhling, J., Lall, P., The Effects of Aging on the Cyclic Stress-strain Behavior and Hysteresis Loop Evolution of Lead Free Solders, Electronic Components and Technology Conference (ECTC), 2011.

Navon, I., Data Assimilation for Numerical Weather Prediction: A Review, Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, Springer-Verlag, 2009.

Noel, S., Correia, S., Alamarguy, D. and Gendre, P., Fretting Behavior of Various Intermetallic Compounds in Electrical Contacts: Influence on Reliability, Wear, Vol. 271, pp. 1515-1523, 2011.

Orchard, M., Vachtsevanos, G., A Particle Filtering Approach for On-Line Failure Prognosis in a Planetary Carrier Plate, International Journal of Fuzzy Logic and Intelligent Systems, Vol. 7, No. 4, pp. 221–227, 2007.

Orchard, M., Vachtsevanos, G., A particle-filtering approach for on-line fault diagnosis and failure prognosis. Transactions of the Institute of Measurement & Control, No. 3, 2009.

Orfanidis, S., Introduction to Signal Processing. Prentice Hall, 1996.

Padalkar, S., Karsai, G., Biegl, C., Sztipanovits, J., Okuda, K., Miyasaka, N., Real-time Fault Diagnostics, IEEE Expert, Vol. 6, No. 3, pp.75–85, 1991.

Prassana, R, Engel, P., Constable, J., An Improved Resistance Method to Determine Lead Force and Lead Fatigue in Twisted Circuit Board, ASME InterPack, 1995.

Pecht, M. Handbook of Electronic Package Design. CRC, 1991.

Pitarresi, J., Caletka, D., Caldwell, R., Smith, D., The "Smeared" Property Technique for the FE Vibration Analysis of Printed Circuit Cards, Journal of Electronic Packaging, Vol. 113, No. 3, pp 250–257, 1991.

Polchow, J., Angadi, S., Jackson, R., Choe, S.,Flowers, G., Lee, B., Zhong, L., A Multi-Physics Finite Element Analysis of Round Pin High Power Connectors, Holm Conference On Electrical Contacts, 2010.

Pullum, L.L., Taylor, B., Darrah, M. Guidance for the Verification and Validation of Neural Networks, Vol. 11. Wiley-IEEE Computer Society, 2007.

Qi, H., Ganesan, S., Wu, J., Pecht, M., Matkowski, P., Felba, J., Effects of Printed Circuit Board Materials on Lead-free Interconnect Durability, International Conference on Polymers and Adhesives in Microelectronics and Photonics, 2005.

Ramakrishnan, A., Pecht, M. A Life Consumption Monitoring Methodology for Electronic Systems IEEE Transactions Components and Packaging Technologies, Vol. 26, No. 3 pp 625–634, 2003.

Rappaport, E., et al., Advances and Challenges at the National Hurricane Center, Weather and Forecasting, vol. 24, no. 2, pp. 395–419, 2009.

Ristic, B., Arulampalam, S., Gordon, N., Beyond the Kalman filter: Particle filters for tracking applications. Artech House Publishers, 2004.

Roebber, P., Bosart, L., The Complex Relationship Between Forecast Skill and Forecast Value: A Real-World Aanalysis. Weather and Forecasting, 11, no. 4 (1996): 544–559.

Roemer, M., Dzakowic, J., Orsagh, R., Byington, C., Vachtsevanos, G., Validation and Verification of Prognostic and Health Management Technologies, IEEE Aerospace Conference, 2005.

Roemmich, D., et. al., The Argo Program : observing the global ocean with profiling floats, Oceanography Society, Vol. 22, No. 2, 34-43, 2009.

Rorgren, R., Tegehall, P., Carlsson, P., Reliability of BGA Packages in an Automotive Environment. Journal of Surface Mount Technology, Vol. 11, pp. 35–44, 1998.

Ruston, B., McNally, A., Explicit handling of surface emission for the exploitation of high spectral resolution infrared satellite sounding radiances from IASI over land and sea-ice, EUMETSAT Satellite Application Facility on Numerical Weather Prediction, February, 2012.

Sandborn, P., Wilkinson, C., A Maintenance Planning and Business Case Development Model for the Application of Prognostics and Health Management (PHM) to Electronic Systems, Microelectronics Reliability, Vol. 47, no. 12, pp. 1889-1901, 2007.

Saha, B., Goebel, K., Uncertainty Management for Diagnostics and Prognostics of Batteries using Bayesian Techniques, Proceedings of the IEEE Aerospace Conference, 2008.

Saha, B., Goebel, K., Modeling Li-ion Battery Capacity Depletion in a Particle Filtering Framework, Proceedings of the Annual Conference of the Prognostics and Health Mngt Society, http://www.phmsociety.org/node/167, San Diego, CA, pp. 1-10, 2009a.

Saha, B., Goebel, K., Prognostics Methods for Battery Health Monitoring Using a Bayesian Framework, , IEEE Transactions on Instrumentation and Measurement, Vol. 58, No. 2, pp. 291-296, 2009b.

Saha, B., Goebel, K., Model Adaptation for Prognostics in a Particle Filtering Framework, International Journal of Prognostics and Health Management , 2011a.

Saha, S., Celaya, J, Vashchenko, V., Mahiuddin, S., Goebel, K., Accelerated Aging with Electrical Overstress and Prognostics for Power MOSFETs, IEEE EnergyTech, 2011b.

Savitzky, A., Golay, M., Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry, Vol. 36, No. 8, pp. 1627–1639, 1964.

Saxena, A., Celaya, J., Saha, B., Saha, S., Goebel, K., Evaluating Algorithm Performance Metrics Tailored for Prognostics, IEEE Aerospace Conference, 2008a.

Saxena, A., Celaya, J., Balaban, E., Goebel, K., Saha, B., Saha, S., Schwabacher, M., Metrics for Evaluating Performance of Prognostic Techniques, Intl. Conf. on Prognostics and Health Management, Denver, Colorado, pp. 1-17, October 2008b.

Saxena, A., Celaya, J., Saha, B., Goebel, K., Evaluating algorithm performance metrics tailored for prognostics. IEEE Aerospace conference, 2009a.

Saxena, A., Celaya, J., Saha, B., Saha, S., Goebel, K., On Applying the Prognostics Performance Metrics, Annual Conference of the PHM Society, vol 1, San Diego, CA, 2009b.

Schwabacher, M., Goebel, K., A Survey of Artificial Intelligence for Prognostics, AAAI Fall Symposium, 2007.

Shang, H., Actuarial Science: Theory and Methodology. World Scientific, 2006.

Simon, D., Simon, D.L., Aircraft Turbofan Engine Health Estimation Using Constrained Kalman Filtering, Journal of Engineering for Gas Turbines and Power, Vol. 127, 2005.

Solomou, S., Economic cycles. Manchester University Press, 1998.

Sornette, D., Davis, A., Kamm, J., Ide, K., Computational Methods in Transport: Verification and Validation, Lecture Notes in Computational Science and Engineering, Springer Berlin Heidelberg, 2008.

Srinivasan, M., Sadhu, S., Ghoshal, T., Ground-Based Radar Tracking of Ballistic Target on Re-entry Phase Using Derivative-Free Filter, IEEE India Conference, 2006.

Steinier, J., Termonia, Y., Deltour, J., Smoothing and Differentiation of Data by Simplified Least Square Procedure, Analytical Chemistry, Vol. 44, No. 11, pp. 1906–1909, 1972.

Stienberg, D., Vibration Analysis for Electronic Equipment, John Wiley and Sons, 2000.

Strickland, S.M. , Hester, J.D., Gowan, A.K., Montgomery, R.K., Geist, D.L., Blanche, J.F., McGuire, G.D., Nash, T.S., Microcoil Spring Interconnects for Ceramic Grid Array Integrated Circuits, NASA Technical Memorandum, 2011.

Suhir, E. Accelerated Life Testing (ALT) in Microelectronics and Photonics: Its Role, Attributes, Challenges, Pitfalls, and Interaction With Qualification Tests, Journal of Electronic Packaging Vol. 124, no. 3,pp. 281–291 2002.

Swanson, D., Spencer, M., Arzoumaniana, S., Prognostic Modeling of Crack Growth in a Tensioned Steel Band, Mechanical Systems and Signal Processing, Vol. 14, No. 5, pp. 789-803, 2000.

Swanson, D., A General Prognostic Tracking Algorithm for Predictive Maintenance, IEEE Aerospace Conference, 2001

Swingler, J., Enhancing Connector Reliability by Using Conducting Polymer Materials to Minimise Contact Fretting, Materials & Design, Vol. 30, No. 10, pp. 3935–3942, 2009.

Swingler, J., Lam, L. and McBride, J. W., Study of Temperature Change and Vibration Induced Fretting on Intrinsically Conducting Polymer Contact Systems, IEEE Transactions On Components and Packaging Technology, Vol. 33, No. 2, pp. 409-415, 2010.

Temple, P., An Introduction to Phase-sensitive Amplifiers: An Inexpensive Student Instrument, American Journal of Physics, Vol. 43, No. 9 pp. 801 – 80, 1973.

Tu, P, Chan, Y., Lai, J., Effect of Intermetallic Compounds on Vibration Fatigue of μBGA Solder Joints, IEEE Transactions on Advanced Packaging, Vol. 24, no. 2, pp. 197-205, 2011.

Valkenburg, M. E. Van. Network Analysis, Prentice Hall, 1955.

Varcic, Aljosa, Kresimir Rupnik, and Sean McGlynn. Selective digital integrator, January 23, 2001. US Patent 6177895.

Verma, V., Gordon, G., Simmons, R., Thrun, S., Particle filters for rover fault diagnosis. IEEE Robotics & Automation Magazine special issue on Human Centered Robotics and Dependability, 2004.

Vetterling, William ,.Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, 1992.

Vichare, N., Pecht, M., Prognostics and Health Management of Electronics, IEEE Transactions on Components and Packaging Technologies, Vol. 29, No. 1, pp. 222– 229, 2006.

Voutilainen, J., Putaala, J., Moilanen, M., Jantunen, H., A Prognostic Method for the Embedded Failure Monitoring of Solder Interconnections with 1149.4 Test Bus Architecture, Microelectronics Journal, vol. 40, no. 7, pp. 1069-1080, Jul. 2009.

Voutilainen, J. Häkkinen, J.,Moilanen, M., Solder Interconnection Failure Time Estimation Based on the Embedded Precursor Behaviour Modelling, Microelectronics Reliability, 2010.

Wang, J., Yonggao, J., Dingzhang, D., Huachun, D., Taifan, Q., Particle Filter Initialization in Non-linear non-Gaussian Radar Target Tracking, Journal of Systems Engineering and Electronics, Vol. 18, No. 3, pp. 491–496, 2007.

Welford, J. Multi-sensor Debris Tracking, IET Seminar on Target Tracking and Data Fusion: Algorithms and Applications, 2008.

Wheeler, A. and A. Ganji, Introduction to Engineering Experimentation, Pearson Education, 2004.

Whiton, R., Smith, P., History of Operational Use of Weather Radar by U.S. Weather Services. Part II: Development of Operational Doppler Weather Radars, Weather and Forecasting, Vol. 13, 1989.

Wilkinson, C., Humphrey, D., Vermeire, B., Houston, J., Prognostic and Health Management for Avionics, IEEE Aerospace Conference, 2004.

Xie, F., Flowers, G. T., Chen, C., Bozack, M. J., Suhling, J., Rickett, B. I., Malucci, R. D., Manlapaz, C., Analysis and Prediction of Vibration-Induced Fretting Motion in a Blade/Receptacle Connector Pair, IEEE Transactions On Components and Packaging Technology, Vol. 32, No. 3, pp. 583-590, 2009.

Yeo, A., Lee, C., Pang, J., Flip Chip Solder Joint Reliability Analysis Using Viscoplastic and Elastic-plastic-creep Constitutive Models, IEEE Transactions on Components and Packaging Technologies, Vol. 29, No. 2, pp355–363, 2006.

Zarchan, P., and H. Musoff, Fundamentals of Kalman Filtering: A Practical Approach, Vol. 190. Progress in Astronautics and Aeronautics, American Institute of Aeronautics and Astronautics (AIAA), 2000.

Zhang, T., McHugh, P., Leen, S., Computational Study on the Effect of Contact Geometry on Fretting Behaviour, Wear, Vol. 271, No. 9–10, pp.1462–1480, 2011.

Zhao, Y., Basaran, C., Cartwright, A., Dishongh, T., Thermomechanical Behavior of Micron Scale Solder Joints Under Dynamic Loads. Mechanics of Materials, Vol. 32, no. 3, pp. 161–173, 2000