

## ABSTRACT

Title of Document: A GENERAL CAUSE BASED METHODOLOGY  
FOR ANALYSIS OF DEPENDENT FAILURES IN  
SYSTEM RISK AND RELIABILITY  
ASSESSMENTS

Andrew N O'Connor  
Doctoral of Philosophy  
Reliability Engineering  
2013  
andrewnoconnor@gmail.com

Directed By: Professor Ali Mosleh  
Reliability Engineering Program  
Mechanical Engineering Department  
University of Maryland

Traditional parametric Common Cause Failure (CCF) models quantify the soft dependencies between component failures through the use of empirical ratio relationships. Furthermore CCF modeling has been essentially restricted to identical components in redundant formations. While this has been advantageous in allowing the prediction of system reliability with little or no data, it has been prohibitive in other applications such as modeling the characteristics of a system design or including the characteristics of failure when assessing the risk significance of a failure or degraded performance event (known as an event assessment).

This dissertation extends the traditional definition of CCF to model soft dependencies between like and non-like components. It does this through the explicit modeling of

soft dependencies between systems (coupling factors) such as sharing a maintenance team or sharing a manufacturer. By modeling the soft dependencies explicitly these relationships can be individually quantified based on the specific design of the system and allows for more accurate event assessment given knowledge of the failure cause.

Since the most data informed model in use is the Alpha Factor Model (AFM), it has been used as the baseline for the proposed solutions. This dissertation analyzes the US Nuclear Regulatory Commission's Common Cause Failure Database event data to determine the suitability of the data and failure taxonomy for use in the proposed cause-based models. Recognizing that CCF events are characterized by full or partial presence of "root cause" and "coupling factor" a refined failure taxonomy is proposed which provides a direct link between the failure cause category and the coupling factors.

This dissertation proposes two CCF models (a) Partial Alpha Factor Model (PAFM) that accounts for the relevant coupling factors based on system design and provide event assessment with knowledge of the failure cause, and (b) General Dependency Model (GDM), which uses Bayesian Network to model the soft dependencies between components. This is done through the introduction of three parameters for each failure cause that relate to component fragility, failure cause rate, and failure cause propagation probability.

A GENERAL CAUSE BASED METHODOLOGY FOR ANALYSIS OF  
DEPENDENT FAILURES IN SYSTEM RISK AND RELIABILITY  
ASSESSMENTS

By

Andrew N. O'Connor.

Dissertation submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
2013

Advisory Committee: Professor Ali Mosleh (Chair)  
Professor Mohammad Modarres  
Associate Professor Michel Cukier  
Associate Professor Jeffrey Herrmann  
Professor Gregory Baecher (Dean's Representative)

© Copyright by  
Andrew N. O'Connor  
2013

[andrewnoconnor@gmail.com](mailto:andrewnoconnor@gmail.com)

## Dedication

*To my wife, Wendy, who has dedicated her life to me, and to my three daughters who remind me why.*

## Acknowledgements

I would like to express my sincere gratitude to the many people who have assisted me in completing this study.

To my wife Wendy. Whilst I live a nocturnal life, she has raised three children, moved our house between two countries and three states and supported me in every way possible, but mostly through the provision of time and food. This was a joint effort. I could not have done this without her. I love you. To our parents; Kerry, Berry, Harry and Debbie. Thanks for the support you have provided our family. We wouldn't be here without you (literally and figuratively).

To my advisor, Professor Ali Mosleh; my first true mentor. Your encouragement, knowledge, experience, critical thinking and positive influence have been invaluable. This study was inspired and driven by your belief. To the staff of the Nuclear Regulatory Commission. Thank you for your invaluable insights and access to the information and data that has enabled this study.

To the small army of reviewers; Thorsten Kels, Glenn Keppel, Peter Stuart, Andrew Kelly, Wendell Fox, Marek Druzdzal, Ben Mailler, Chris Jackson and Patrick O'Neill. Thank you. To my business partners Ben Mailler and Chris Jackson, thank you for giving me the time and opportunity to finish this dissertation.

To all those I've missed. Now that I have time, I owe you a beer.

# Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Tables.....	x
List of Figures.....	xiii
Chapter 1: Introduction.....	1
1.1. Introduction.....	1
1.2. Significance.....	3
1.3. Objectives of the research.....	6
1.3.1. Unified CCF Definition.....	8
1.3.2. Propose Failure Data Taxonomy.....	9
1.3.3. Propose a cause-based data-informed CCF model.....	9
1.4. Structure of the dissertation.....	14
Chapter 2: Current CCF Analysis Methodology and its Limitations.....	16
2.1. Introduction.....	16
2.2. Describing CCFs.....	18
2.3. Problem Definition and System Modeling.....	19
2.3.1. Example 1: Two Train EDG System.....	19
2.3.2. Example 2: Two Train EDG and Pump System.....	21
2.4. Preliminary Analysis of CCF Vulnerabilities.....	24
2.4.1. Qualitative Screening.....	24
2.4.2. Quantitative Screening.....	27
2.5. Detailed Qualitative Analysis.....	28
2.6. Detailed Qualitative Analysis.....	28
2.6.1. Identification of Common Cause Basic Events (CCBEs).....	28
2.6.2. Incorporate CCBEs into fault trees.....	29
2.6.3. Parametric representation of CCBEs.....	33
2.6.4. Alpha Factor Model Parameterization.....	36
2.7. Data Analysis and Parameter Estimation.....	37
2.7.1. Parameter Estimation – Impact Vectors.....	37
2.7.2. Parameter Estimation - Alpha Factor Model.....	44
2.8. System Quantification and Results Interpretation.....	46
2.8.1. System unavailability quantification.....	46
2.8.2. Results Evaluation and Sensitivity Analysis.....	47
2.8.3. Reporting.....	48
2.9. Asymmetrical Components.....	48
2.10. Impact Vector Mapping.....	51
2.10.1. Mapping Up Independent Events.....	51
2.10.2. Mapping Up :Lethal Shocks.....	52
2.10.3. Mapping Up Non-Lethal Shocks.....	52

2.11.	Event Assessment .....	56
2.11.1.	Event Assessment Using AFM .....	56
2.12.	Current Issues in CCF Modeling Specific to Event Assessment .....	59
2.13.	Mission Time .....	60
2.14.	Summary of Issues .....	62
Chapter 3:	Definition of Common Cause Failure .....	65
3.1.	Introduction .....	65
3.2.	Test Cases .....	66
3.3.	Literature Review .....	69
3.4.	Definition Discussion .....	73
3.4.1.	Multiplicity of Failure .....	73
3.4.2.	Simultaneity .....	73
3.4.3.	Functional Failure and Physical Failure (System Boundary) .....	75
3.4.4.	Independent and Common Cause Failures .....	75
3.4.5.	Explicit and Implicit Dependencies .....	77
3.4.6.	Redundant Components .....	79
3.4.7.	Identical Components .....	80
3.5.	Proposed CCF Definition .....	80
Chapter 4:	Common Cause Failure Database Taxonomy .....	83
4.1.	Literature Review .....	84
4.2.	Analysis of NRC Common Cause Failure Database .....	88
4.2.1.	CCFDB Classification System .....	89
4.2.2.	Analysis of Observed CCF Events .....	91
4.3.	Failure Cause and Coupling Factor Taxonomy Proposal .....	94
4.4.	Summary .....	96
Chapter 5:	Existing CCF models .....	97
5.1.	Introduction .....	97
5.2.	Direct Estimates .....	98
5.2.1.	Direct Assessment (Qualitative) .....	98
5.2.2.	Basic Parameter Model .....	99
5.3.	Ratio Models .....	101
5.3.1.	Beta Factor Model .....	102
1.1.1.	Partial Beta Factor (PBF) Model .....	104
1.1.2.	Alpha Factor Model (AFM) .....	105
5.4.	Shock Models .....	107
5.4.1.	Binomial Failure Rate Model .....	109
5.4.2.	Binomial with Lethal Shocks .....	111
5.5.	Interference Models .....	113
5.5.1.	Common Load Model .....	115
5.6.	Other Models .....	116
5.6.1.	Reliability Cut Off Method .....	116
5.6.2.	Unified Partial Method .....	117
5.6.3.	Influence Diagram Model (Zitrou 2006a) .....	117



5.6.4.	CCF model for Event Assessment (Kelly et al. 2011).....	120
5.7.	Model Comparison.....	121
Chapter 6:	Partial Alpha Factor Model (PAFM) .....	124
6.1.	Introduction.....	124
6.1.1.	Motivation for PAFM .....	124
6.1.2.	Chapter Scope .....	125
6.1.3.	Examples.....	126
6.2.	Model Overview .....	126
6.3.	Parameter Description.....	128
6.4.	Parameter Estimation .....	128
6.4.1.	Partial Alpha Factor .....	129
6.4.2.	Gamma Factor.....	132
6.4.3.	Assessed Alpha Factor .....	134
6.5.	Parameter Quantification .....	135
6.5.1.	Using Impact Vectors .....	135
6.5.2.	Using Generic Data Sources .....	136
6.5.3.	Informative Prior Distributions.....	142
6.5.4.	Non-informative Prior Distributions.....	144
6.5.5.	Using Alpha Factors as a Constraint.....	146
6.6.	PAFM in System Analysis.....	148
6.6.1.	Qualitative Analysis.....	149
6.6.2.	Identification of Common Cause Basic Events .....	150
6.6.3.	Incorporate into Fault Tree .....	151
6.6.4.	Parametric representation of CCBEs .....	153
6.6.5.	Partial Alpha Factor Model Parameterization.....	154
6.6.6.	Parameter Estimation – Impact Vectors .....	156
6.6.7.	Parameter Estimation – Partial Alpha Factor Model .....	157
6.6.8.	System Quantification and Results Interpretation. ....	158
6.7.	PAFM in Event Assessment .....	160
6.7.1.	Knowledge of Failure .....	161
6.7.2.	Knowledge of Failure Cause.....	163
6.8.	Data Collection Requirements .....	167
6.8.1.	Desirable Data Collection Attributes .....	167
6.8.2.	When a one to one relationship between cause and coupling factor does not exist. ....	170
6.9.	Model Assessment .....	172
6.9.1.	Model Advantages .....	173
6.9.2.	PAFM Limitations .....	173
6.9.3.	Compare Against Model Criteria.....	174
Chapter 7:	General Dependency Model .....	177
7.1.	Introduction.....	177
7.1.1.	Motivation.....	177
7.1.2.	Chapter Scope .....	180

7.1.3.	Examples.....	181
7.2.	Model Structure .....	181
7.2.1.	Component Failure Probability.....	181
7.2.2.	Component Dependency.....	183
7.2.3.	Propagation of Cause Condition.....	184
7.2.4.	Parameter Description.....	188
7.3.	Bayesian Networks .....	196
7.3.1.	Bayesian Network Components.....	198
7.3.2.	Bayesian Network Features .....	200
7.4.	GDM Bayesian Network Structure.....	204
7.4.1.	Component Failure Node.....	204
7.4.2.	Cause Condition Nodes.....	205
7.4.3.	Graphical Representation of GDM .....	216
7.5.	Parameter Estimation .....	221
7.5.1.	GDM Relationship to $Qt, i$ .....	221
7.5.2.	GDM Relationship to $\alpha 2, i$ .....	223
7.5.3.	Estimation Using Observed Data.....	226
7.5.4.	Parameter Uncertainty Calculation.....	229
7.6.	Parameter Quantification .....	229
7.6.1.	Direct Estimates .....	229
7.6.2.	Using Impact Vectors and Causes .....	232
7.6.3.	Coupling Factor Strength Assumption.....	234
7.6.4.	Existing Parametric Model Estimate .....	234
7.6.5.	Prior Distributions.....	236
7.7.	GDM in System Analysis .....	237
7.7.1.	Qualitative Analysis.....	239
7.7.2.	Create GDM Structure with Common Characteristics .....	241
7.7.3.	Identify Constraints from Observable Quantities .....	245
7.7.4.	Estimate Parameters within Constraints. ....	247
7.7.5.	Calculate Model.....	253
7.8.	GDM in Event Assessment.....	255
7.8.1.	Knowledge of Failure .....	256
7.8.2.	Knowledge of Failure Cause.....	258
7.8.3.	Uncertain Knowledge of Failure Cause.....	262
7.9.	Data Collection Requirements .....	266
7.10.	Model Assessment .....	267
7.10.1.	Model Advantages .....	267
7.10.2.	GDM Limitations.....	269
7.10.3.	Compare Against Model Criteria.....	270
7.11.	Extensions and Future Development of GDM.....	273
7.11.1.	Scalability .....	273
7.11.2.	Lethal Shocks.....	276
7.11.3.	Consistency of Asymmetrical Components.....	278

7.11.4.	Uncertain Evidence In Event Assessment .....	279
7.11.5.	Failure Taxonomy Development .....	281
7.11.6.	Prior Estimations.....	281
7.11.7.	Unbiased Estimators .....	281
7.11.8.	Integration of Existing Parametric Models .....	282
7.11.9.	Software and Procedure Development.....	282
Chapter 8:	Conclusion .....	284
8.1.	Introduction.....	284
8.2.	Review of research objectives and goals .....	284
8.3.	Common Cause Failure Definition .....	286
8.4.	CCF Failure Taxonomy .....	289
8.5.	CCF Model – Partial Alpha Factor Model.....	289
8.5.1.	Overview.....	289
8.5.2.	Advantages.....	291
8.5.3.	Limitations .....	292
8.6.	CCF Model - General Dependency Model .....	293
8.6.1.	Overview.....	293
8.6.2.	Advantages.....	294
8.6.3.	Limitations .....	295
8.6.4.	Future Work for GDM .....	296
8.7.	Future Research .....	298
8.7.1.	Time Relationship of CCF (Repair Time, Mission Time, Aging)....	299
8.7.2.	Existing Parametric Model Estimate .....	301
8.7.3.	Quantitative Defense Modeling .....	301
8.8.	Conclusion .....	302
Appendices 1:	Literature Review of CCF Models.....	304
1.1.	Introduction.....	304
1.2.	Direct Estimates .....	305
1.2.1.	Direct Assessment (Qualitative) .....	305
1.2.2.	Basic Parameter Model .....	306
1.3.	Ratio Models.....	308
1.3.1.	Beta Factor Model.....	310
1.3.2.	C-Factor Model.....	312
1.3.3.	Multiple Beta Factor (MBF) Model.....	312
1.3.4.	Multiple Dependent Failure Fraction (MDFF) .....	314
1.3.5.	Partial Beta Factor (PBF) Model .....	314
1.3.6.	Multiple Greek Letter Model .....	316
1.3.7.	Alpha Factor Model (AFM).....	318
1.4.	Shock Models.....	319
1.4.1.	Binomial Failure Rate Model.....	322
1.4.2.	Binomial with Lethal Shocks.....	324
1.4.3.	Rho Distribution Models.....	325
1.4.4.	Multinomial Failure Rate Model .....	327

1.4.5.	Stochastic Reliability Analysis Models .....	327
1.4.6.	Trinomial Failure Rate Model .....	328
1.4.7.	Multi-Class Binomial Failure Rate Model.....	328
1.4.8.	The Coupling Model.....	330
1.4.9.	Bayes Testing and Estimation BFR Model.....	331
1.5.	Interference Models .....	332
1.5.1.	Common Load Model.....	334
1.5.2.	Inverse Stress-Strength Interference Model (ISSI).....	336
1.5.3.	Harris Model .....	336
1.5.4.	Knowledge Based Multi-dimension CCF Model (KBMD).....	337
1.6.	Other Models .....	338
1.6.1.	Square Root Bounding Method .....	338
1.6.2.	Implicit Method .....	338
1.6.3.	Reliability Cut Off Method.....	339
1.6.4.	Unified Partial Method .....	340
1.6.5.	Influence Diagram Model (Zitrou 2006a).....	340
1.6.6.	CCF model for Event Assessment (Kelly et al. 2011).....	342
1.6.7.	Physics-Based CCF.....	343
1.7.	Model Comparison.....	344
Appendices 2: Detailed Description of Existing Failure Data Taxonomy .....		348
2.1.	Introduction.....	348
2.2.	Failure Causes:.....	348
2.3.	Coupling Factor Definitions .....	350
2.3.1.	Environmental Based.....	350
2.3.2.	Design Based .....	351
2.3.3.	Quality Based.....	351
2.3.4.	Maintenance Based.....	352
2.3.5.	Operation Based.....	352
Appendices 3: Calculation of Mutually Exclusive Nodes Using Control Node... 354		
3.1.	The Bayesian Network.....	354
3.2.	Calculate Control Node States .....	356
3.3.	Summary .....	359
Appendices 4: Calculation of Event Assessment for GDM Example 1 .....		360
4.1.	The Bayesian Network.....	360
4.2.	Calculate Evidence Propagation .....	362
4.3.	Summary .....	368
Notation.....		369
Glossary .....		373
Abbreviations .....		376
Bibliography .....		377

## List of Tables

Table 1: Comparison of results from modeling EDG power system failure .....	5
Table 2: Criterion for CCF Model Assessment .....	12
Table 3: Qualitative dependency assessment for example 1 .....	25
Table 4: Qualitative dependency assessment for example 2 .....	26
Table 5: CCBE for example 1.....	29
Table 6: CCBE for example 2.....	29
Table 7: Example Failure Data for Emergency Diesel Generator .....	42
Table 8: Summary of Impact Vectors for EDG by Cause .....	42
Table 9: Example Failure Data for Pump .....	43
Table 10: Summary of Impact Vectors for Pump by Cause .....	43
Table 12: Comparison of 4 train and 2 train system basic events .....	54
Table 13: Cut Sets for Example 2 in event assessment .....	59
Table 14: Comparison of CCF definitions for each test case .....	81
Table 15: Comparison of NUREF/CR-6268 Failure Cause Classification and the CCFDB .....	90
Table 16: Comparison of NUREF/CR-6268 Coupling Factor Classification and the CCFDB .....	91
Table 17: Comparison of failure cause and coupling factor for observed CCF events in the CCFDB.....	92
Table 18: Proposed Cause and Coupling Factor Taxonomy .....	95
Table 19: Assessment of previous CCF Models.....	122
Table 20: Comparison of weighting factor strengths.....	142
Table 21: Qualitative dependency assessment for example 1 .....	149
Table 22: Qualitative dependency assessment for example 2 .....	149
Table 23: CCBE for example 1.....	151
Table 24: CCBE for example 2.....	151
Table 25: Cut Sets for Example 2 in event assessment .....	162

Table 26: Event Assessment for Example 1 with different failure causes .....	165
Table 27: Cut Sets for Example 2 in event assessment .....	166
Table 28: Event Assessment for Example 2 with different failure causes .....	167
Table 29: Demonstration of impact vectors changing for target system based on assumption of CCCG size for observed single failure.....	169
Table 30: Example data when a failure cause can propagate through multiple coupling factors.....	171
Table 31: Assessment of the PAFM compared to previous CCF models.....	175
Table 32: Comparison of probability and rate metrics .....	194
Table 33: CPT for Control Node .....	205
Table 34: CPT for Control Node .....	212
Table 35: CPT for Virtual Evidence Node .....	214
Table 36: CPT for Local Cause Condition Node.....	215
Table 37: CPT for Common Cause Condition <b><i>Xi</i></b> .....	216
Table 38: CPT for Independent Cause Condition <b><i>Ii</i></b> .....	216
Table 39: CPT for No Cause Condition <b><i>Ni</i></b> .....	216
Table 40: Qualitative dependency assessment for example 1 .....	217
Table 41: Qualitative dependency assessment for example 1 .....	239
Table 42: Qualitative dependency assessment for example 2 .....	240
Table 43: Estimation of <b><i>Qt, i</i></b> for EDG .....	246
Table 44: Estimation of <b><i>Qt, i</i></b> for Pump .....	247
Table 45: GDM Parameter Estimates for example 1 EDG.....	249
Table 46: GDM Parameter Estimates for example 2 .....	252
Table 47: Event Assessment for Example 1 with different failure causes .....	260
Table 48: Event Assessment for Example 2 with different failure causes .....	262
Table 49: Assessment of the GDM compared to previous CCF models .....	271
Table 50: Comparison of CCF model features .....	346
Table 51: CPT for Common Cause Condition <b><i>Xi</i></b> .....	355
Table 52: CPT for Independent Cause Condition <b><i>Ii</i></b> .....	355

Table 53: CPT for No Cause Condition <b><i>Ni</i></b> .....	355
Table 54: CPT for Control Node .....	355
Table 55: CPT for Virtual Evidence Node .....	356
Table 56: CPT for Local Cause Condition Node.....	356
Table 57: CPT for Common Cause Condition <b><i>Xi</i></b> .....	361
Table 58: CPT for Independent Cause Condition <b><i>Ii</i></b> .....	361
Table 59: CPT for No Cause Condition <b><i>Ni</i></b> .....	361
Table 60: CPT for Control Node .....	361
Table 61: CPT for Virtual Evidence Node .....	362
Table 62: CPT for Local Cause Condition Node.....	362
Table 63: Cause Condition Node <b><i>CA</i></b> .....	363
Table 64: Virtual Evidence Node <b><i>VA</i></b> .....	363
Table 65: Control Node <b><i>LA</i></b> : .....	364
Table 66: No Cause Condition Node <b><i>NA</i></b> .....	364
Table 67: Independent Cause Condition Node <b><i>IA</i></b> .....	365
Table 68: Common Cause Condition Node <b><i>Xx</i></b> .....	365
Table 69: Control Node <b><i>LB</i></b> .....	366
Table 70: Cause Condition Node <b><i>CB</i></b> .....	367
Table 71: Second Component Node ( <b><i>B</i></b> ).....	367

## List of Figures

Figure 1: Fault Tree of redundant generator system not considering CCF.....	4
Figure 2: Fault Tree of redundant generator system considering CCF.....	4
Figure 3: Qualitative and quantitative features of proposed CCF models (Lindberg 2007) .....	11
Figure 4: Procedural framework for common cause failure analysis (Mosleh et al. 1998, p.10) .....	17
Figure 5: Conceptual diagram for example 1- Two train EDG system .....	20
Figure 6: Reliability block diagram for example 1- Two train EDG system.....	20
Figure 7: Fault tree for example 1 – Two train EDG system.....	20
Figure 8: Conceptual diagram for example 2- Two EDGs and three pump system ...	22
Figure 9: Reliability block diagram for example 2- Two EDGs and three pump system .....	23
Figure 10: Fault tree for example 2 – Two train EDG and pump system.....	24
Figure 11: Conversion of Basic Events to CCBE.....	30
Figure 12: Fault tree for example 1 with CCBEs .....	30
Figure 13: Fault tree for example 2 with CCBEs .....	31
Figure 14: Basic Event Fault Tree for A.....	32
Figure 15: Representation of alpha factor parameters to failure and demand data-set45	
Figure 16: Two Train EDG Fault Tree with Failure Modes.....	49
Figure 17: : The probability distribution for the number of failures at the end of a mission (O).....	57
Figure 18: The probability distribution for the number of failures at the end of a mission time (O), after a failure has been observed. ....	57
Figure 19: Recording of CCF Events into CCFDB .....	62
Figure 20: Test Case 1 scenario .....	67
Figure 21: Test Case 2 scenario .....	67
Figure 22: Test Case 3 scenario .....	68



Figure 23: RDRC Diagram(Lindberg 2007, p.31).....	88
Figure 24: Zitrou General Influence Diagram Structure.....	118
Figure 25: Bayesian network representing more general situation of multiple failure mechanisms and causes in a CCCG of two EDGs (Kelly et al. 2011, p.6) .....	121
Figure 26: Failure Event Conditionalisation.....	125
Figure 27: Reliability block diagram for example 1- Two train EDG system.....	148
Figure 28: RBD for example 2, two EDG three pump system. ....	148
Figure 29: Fault tree for example 1 with CCBEs .....	152
Figure 30: Fault tree for example 2 with CCBEs .....	152
Figure 31: GDM Basic Events .....	183
Figure 32: GDM Coupling Components.....	184
Figure 33: Conceptual propagation of cause condition through coupling factor.....	186
Figure 34: Conceptual construction of the GDM model .....	188
Figure 35: Bayesian Network for Burglary Example .....	198
Figure 36: (A) Causal chain (B) common cause (C) common effect .....	201
Figure 37: Marginal Distributions for burglary example.....	202
Figure 38: (A) Diagnostic Reasoning (B) Predictive Reasoning (C) Explaining Away .....	203
Figure 39: GDM Basic Events.....	204
Figure 40: Conceptual cause condition modeling.....	206
Figure 41: Problem with propagating the common cause condition .....	208
Figure 42: Problem with propagating the common cause condition .....	208
Figure 43: Problem with propagating the common cause condition .....	209
Figure 44: Structure of mutually exclusive Bayesian network.....	211
Figure 45: Structure of mutually exclusive Bayesian network with VE .....	213
Figure 46: Reliability block diagram for example 1- Two train EDG system.....	217
Figure 47: Example GDM Bayesian Network structure for example 1.....	218
Figure 48: Nodes from the GDM Model which may be combined for visual representation .....	220

Figure 49: Compact representation of GDM .....	220
Figure 50: GDM Structure for Example 1 .....	242
Figure 51: GDM Structure for Example 2 .....	243
Figure 52: GDM Structure for Example 2 analyst interface.....	244
Figure 53: GDM system analysis results for example 1 .....	254
Figure 54: GDM system analysis results for example 2.....	255
Figure 55: Event assessment for component A failing using GDM.....	256
Figure 56: Event assessment for component P1 failing using GDM.....	258
Figure 57: Event assessment for component A failing from cause MH using GDM.....	259
Figure 58: Event assessment for component P1 failing due to EE using GDM.....	260
Figure 59: Event assessment for component P1 failing due to IP using GDM .....	261
Figure 60: Event assessment for component Pump 1 prior to applying virtual evidence .....	265
Figure 61: Event assessment for component EDG 1 after applying virtual evidence .....	266
Figure 62: GDM with intermediate nodes .....	274
Figure 63: GDM with parent nodes .....	276
Figure 64: Recording of Failures into CCFDB with consideration for defenses.....	302
Figure 65: Zitrou General Influence Diagram Structure.....	341
Figure 66: Bayesian network representing more general situation of multiple failure mechanisms and causes in a CCCG of two EDGs (Kelly et al. 2011, p.6) .....	343
Figure 67: Classification of CCF Models .....	345
Figure 68: Structure of mutually exclusive Bayesian network with VE .....	354
Figure 69: Structure of mutually exclusive Bayesian network with VE .....	361

# Chapter 1: Introduction

## *1.1. Introduction*

Recent events at the Japanese Fukushima Nuclear Power Plant have reminded us of the importance of protecting safety critical systems against failure causes which can overcome multiple levels of redundancy.

Common Cause Failures (CCF) are ‘simultaneous’ failures of a number of components due to a common event. This phenomenon is usually caused by soft dependencies<sup>1</sup> and can dominate system failure probabilities. These types of failures have the ability to cut through multiple layers of redundancy and cause unforeseen coincidental events that will put safety critical systems in jeopardy.

A Probabilistic Risk Assessment (PRA) studies and quantifies such risks. If CCF modeling is not included within the PRA, the system reliability model can result in a gross overestimation of system safety and reliability.

Over the years, PRA has expanded as a management and decision tool beyond the simple quantification of system failure probabilities. In the nuclear industry PRAs are

---

<sup>1</sup> ‘Soft dependencies’ are dependencies with a probabilistic relationship.

increasingly used to support the following activities (Bates 1995):

- decisions on safety and performance improvement,
- evaluation of proposed modifications,
- assessment of new designs, and
- event assessment and significance determination.

Whilst the current models allow for the quantification of soft dependencies at a system level, they are inadequate for providing further insight into the causes of CCF and informing options for protecting against this phenomenon. It has become evident that the commonly accepted CCF modeling methodology (Mosleh et al. 1998) and corresponding tools need to be enhanced to meet these PRA activities.

Current CCF models can be broadly classified in two ways: qualitative or quantitative (Lindberg 2007). However both approaches have inherent weaknesses. Quantitative CCF models estimate CCF parameters using historic event data but do not currently model the system features needed to conduct these extended PRA activities, such as failure causes, coupling factors and defenses specific to the target system. Qualitative CCF models include the ability to account for the target system features, but rely on expert opinion for quantification and cannot incorporate historic data.

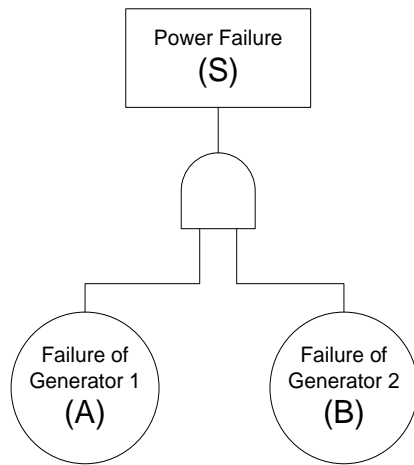
A need exists for a data-informed causal CCF model which can provide both a qualitative investigation of a system while allowing evidence-informed modeling of

CCF events. This thesis will propose such a model and develop the supporting analytical processes and data taxonomy to ensure it can be practically implemented.

### ***1.2. Significance***

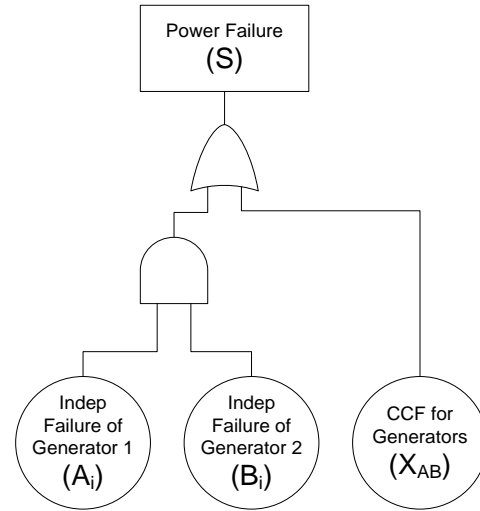
CCFs are usually much less frequent than the independent failures of the components involved. However, studies have shown that CCF events may still contribute between 20% - 80% of the unavailability of safety systems within nuclear reactors (Werner 1994). This is because, despite CCF events being infrequent, they can overcome all levels of redundancy and fail a complete system.

Excluding CCF when modeling system reliability can result in a gross underestimation of failure probability. For example, consider a system which is required to provide power to a safety critical item. The system consists of two backup generators in parallel. Only one generator is required to power the safety critical item, with one redundant generator in case the other fails. Figure 1 shows a fault tree excluding CCF modeling, and figure 2 shows a fault tree including CCF modeling.



$$P(S) = P(A)P(B)$$

**Figure 1: Fault Tree of redundant generator system not considering CCF**



$$P(S) = P(X_{AB}) + P(A_i)P(B_i) - P(X_{AB})P(A_i)P(B_i)$$

**Figure 2: Fault Tree of redundant generator system considering CCF**

An Emergency Diesel Generator (EDG) may have a probability of failure-to-start of  $P(A) = P(B) = 4.9E-3$ . (Vesely et al. 1994). This provides a system failure probability, without consideration for CCF, of:

$$\begin{aligned} P(S) &= P(A)P(B) \\ &= 2.4e-5 \end{aligned}$$

The probability of an EDG failing due to common cause failure may be<sup>2</sup>  $P(X_{AB}) =$

---

<sup>2</sup> Calculated using a  $\beta$  factor of 0.0316 (Wierman et al. 2007, p.78)

1.55E-4. Therefore the probability of an EDG failing due to an independent cause is  $P(A_i) = P(A) - P(X_{AB}) = P(B_i) = 4.745E-3$ . Using the CCF modeling shown in Figure 2 gives the following system failure probability:

$$\begin{aligned}
 P(S) &= P(X_{AB}) + P(A_i)P(B_i) - P(X_{AB})P(A_i)P(B_i) \\
 &= 1.55e-4 + (4.745e-3)^2 - (1.55e-4)(4.745e-3)^2 \\
 &= 1.77e-4
 \end{aligned}$$

Table 1 shows a comparison of system modeling quantification with and without modeling CCF.

**Table 1: Comparison of results from modeling EDG power system failure**

	Probability of System Failure	Expected number of failures
Model excluding CCF	2.4e-5	1 in 41,649 demands
Model including CCF	1.77e-4	1 in 5,638 demands
Factor of difference	7.4	7.4

As demonstrated in this simple example, omitting CCF within the PRA results in a predicted failure frequency 7.4 times more optimistic than is the case when CCF modeling is included.

While this example has shown the importance of including CCF models within a system model, it also shows that in order to conduct more advanced analysis (such as prioritizing system upgrades or post event assessments of system safety), the CCF model needs to provide quantitative insights why CCF occur. Without such insight, management of and protection from CCF becomes very difficult.

Very recently we were reminded of this field of study's significance when a Japanese nuclear plant struggled to cool its reactor core following a major earthquake and tsunami in March 2011. The earthquake provided a single event which would cause the near simultaneous failure of the external power, the emergency power supply and the cooling system reticulation. (Yamaguchi & Donn 2011) The combination of these sub-systems failing without common cause would be extremely unlikely. However, their failures were inter-dependent on vide a single coupling factor of being in the same location<sup>3</sup>.

### ***1.3. Objectives of the research***

The goal of this research is to develop a comprehensive CCF analysis methodology to enhance:

- **CCF analysis in PRA of operating systems.** This form of analysis may be undertaken when the system design is known and failure data are available. While data on common cause failures may not be abundant, the independent failure rates of components are likely to be well known. Furthermore, softer contributors to the system (such as organizational factors) may be incorporated.

---

<sup>3</sup> It should be noted that natural phenomena such as the earthquake and tsunami are explicitly modelled within nuclear PRAs and therefore would not have been included within the CCF modelling scope. However, in non-nuclear PRAs this type of event is likely to have been included within the CCF model scope. A discussion on defining CCF in relation to CCF modelling is provided in chapter 3.



Specific extensions of PRA for operating systems include the ability to provide trade off analysis between different improvement activities.

- **CCF analysis in PRA of systems in design.** This form of analysis will be characterized by a lack of system specific data. Accordingly, it will rely on incorporating evidence from many sources such as generic data, specific life tests and expert opinion. The CCF model may need to have a high level of detail to accurately capture system and component dependencies and influence design decisions. Extensions of the PRA activity during system design include the ability to quantify the affect different design options will have on the system failure probability.
- **CCF analysis in Event Assessments.** This form of analysis will involve retrospective assessments intended to estimate the risk significance of known deficiencies within a system. The model will need to be detailed enough to ensure that the characteristics of the event being assessed can be accurately included within the model, enabling the update of remaining variables to reflect this condition.

In order to support these applications, this dissertation will focus on the following research objectives:

- To propose a unified understanding of the definition and scope of CCFs.
- To propose a failure data taxonomy consistent with the unified definition of CCFs which can enable cause based CCF models.

- To propose a cause based data-informed CCF model.
- To propose a comprehensive and scalable analysis process.

### *1.3.1. Unified CCF Definition*

Since CCF began to be recognized as a phenomenon requiring special treatment, there has not been consensus over its definition. This is primarily because the definition of a CCF and the scope of CCF modeling within a particular system may be different. Modeling CCFs is only used to cover events which are not explicitly included within the wider PRA model. This means that the quantification of CCF can change significantly between PRAs. This issue is a product of conflicting CCF definitions and has limited a unified advancement of the discipline.

An additional problem stems from the definition of a singular failure within a Common Cause Component Group (CCCG). This failure could be considered an ‘independent failure’ which will only manifest itself in single component failure, regardless of the CCCG size. Alternately, it could be considered a CCF failure with potential to cause other components to fail. Different interpretations of this failure type have prompted a myriad of different CCF models.

This dissertation aims to define CCF such that it unifies the observable phenomenon and scope of CCF modeling. Furthermore, this definition aims to make clear the difference between a single and independent failure within a CCF context.

### ***1.3.2. Propose Failure Data Taxonomy***

The treatment of CCF is complicated by the fact that CCF modeling aims to incorporate ‘known unknowns’. We know that certain failures will occur, thereby affecting multiple components as propagated through soft dependencies. However, we typically don’t know the nature of these failures (otherwise they would have modeled explicitly). This makes a consensus on CCF definitions difficult to achieve.

It is for this reason that traditional CCF models have focused on empirical relationships which allow the quantification of PRA models without a detailed definition of failure causes and coupling mechanisms. The symptom of this approach is that very little insight can be provided into the nature of CCFs.

This dissertation will propose the data requirements for the Common Cause Failure Databases (CCFDB) such that there are consistent definitions of failure causes and coupling factors. These definitions will enable data to be recorded in support the CCF models proposed herein.

### ***1.3.3. Propose a cause-based data-informed CCF model***

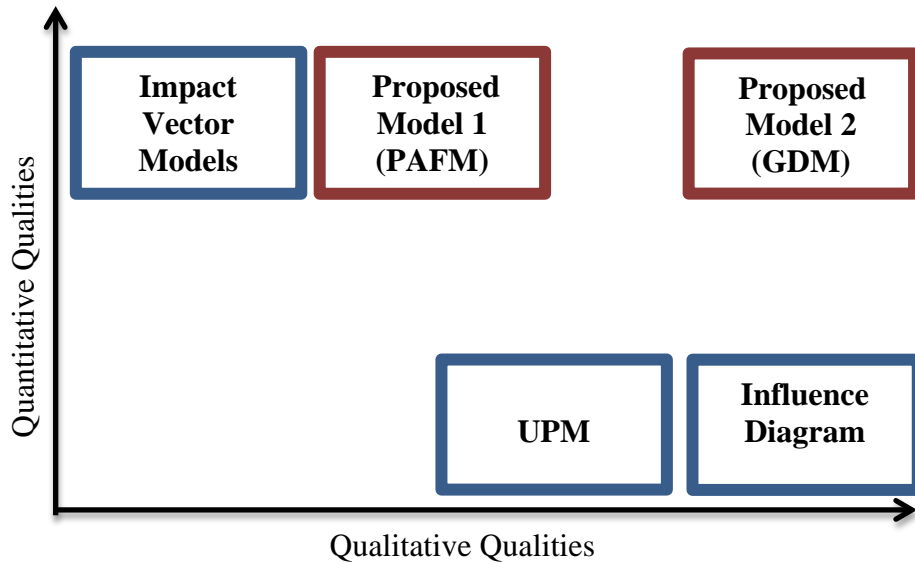
Since the 1975, over 30 CCF models have been proposed, each addressing the particular concerns of their designer. Only a few have gained widespread use due to their support from data collection databases and their simplicity. The US Nuclear Regulatory Commission (NRC) adopted the data-informed parametric models of Alpha

Factor Model and the Multiple Greek Letter models (Mosleh et al. 1988) which provides a ‘black box’ analysis approach using impact vectors for data with little causal information. The UK nuclear industry has adopted the Unified Partial Method (UPM) (Brand & Gabbot 1993) which is a model that assesses system CCF susceptibility, providing an estimate based on expert opinion. UPM is a causal based model providing an excellent qualitative assessment of the system but cannot incorporate CCF event data to improve CCF parameter estimates.

This dissertation will propose two models that can provide both a qualitative investigation of a system while producing evidence-informed system specific estimates of CCF probabilities.

- **Partial Alpha Factor Model (PAFM).** The PAFM is an extension of the current Alpha Factor Model (AFM) that enables more advanced CCF analysis, but minimizes changes to the AFM methodology. This allows backward compatibility.
- **General Dependency Model (GDM).** The GDM achieves the desired level of CCF modeling features without being constrained by current methods. Despite being ‘feature rich,’ it too minimizes complexity for the analyst and remains scalable.

Figure 3 shows a summary of the gap in CCF modeling against quantitative and qualitative features.



**Figure 3: Qualitative and quantitative features of proposed CCF models (Lindberg 2007)**

To assess existing models, a list of desirable CCF model attributes has been developed and is contained in Table 2. These attributes are derived from the features required of a model to support likely PRAs activities that support decisions on new designs, improvements to existing designs and event assessments. These criteria will be also be used to assess the models proposed in this dissertation.

The model criteria covers the following broad aims:

- *Describes System Features.* It is desirable for the model to explicitly account for specific features of the target system in order to increase accuracy. The modeling of specific causes, coupling factors and defenses permits an improved

analysis outcome. This feature set mostly accounts for the qualitative features of the model.

- *CCCG Characteristics.* This list assesses model accuracy by examining the assumptions made in creating the CCCGs. This feature set also contributes to the qualitative features of the model.
- *Event Assessment Capabilities.* This list summarizes the ability for the model to support event assessment activities.
- *Parameter Estimation.* This activity assesses how parameters may be estimated including their ability to use impact vectors, sources from other CCCG sizes and incorporate generic and target system data. This feature set mostly assesses the quantitative features of the model.
- *Uncertainty for Parameter Estimation.* This metric describes the types of uncertainty that can be explicitly accounted for during the estimation of the model parameters.
- *Usability and Cultural Considerations.* Despite some very capable CCF models, the PRA community has been reluctant to adopt methods that require significant time or complexity investment from the analyst. Therefore these considerations describe the usability of the model.

**Table 2: Criterion for CCF Model Assessment**

<b>Feature Description</b>
<b>Explicitly Models System Features</b>
Models failure cause
Models failure cause defense
Models coupling factor
Models coupling factor defense
Models deeper causal levels
Models cause condition / shock
Models multiplicity of failures within CCCG
Models includes consideration for rectification period
<b>Common Cause Component Grouping Characteristics</b>
Model non-symmetrical but similar components within the same CCCG
Model different components within the same CCCG
A component can be part of many CCCGs
No limit to CCCG size
Model different failure multiplicities within the CCCG ( $k$ failures in $n$ )
<b>Event Assessment Capabilities</b>
Event Assessment with knowledge of a failed component
Event Assessment with knowledge of failure cause
Uncertain Evidence - Partial Failures
Uncertain Evidence- Virtual evidence of cause
<b>Parameter Estimation</b>
Impact Vector Method (including method for incorporating uncertainty)
Expert estimations (in absence of any data)
Account for reliability growth (discount previous failures)
Update parameters with new evidence
Incorporate evidence from different sized CCCGs
Account for CCF which occurred in a different mission time
Account for CCF data which has artificial separation in time due to demands being separate.
Use system specific failure rate data combined with generic model parameter

<b>Uncertainty Characteristics for Parameter Estimation</b>
Does not require distinguish between independent and single CCF failures
Failures outside the mission period
Uncertainty of shared cause
Uncertainty of coupling factor
Uncertainty in intervals due to staggered testing
Partial failures and component degradation
<b>Usability and Cultural Considerations</b>
Backward compatible to Alpha Factor Model parameters
The time investment required to implement the model is no more than the alpha factor model.
Automatic parameter estimation is possible from the CCFDB/RADs

#### ***1.4. Structure of the dissertation***

This dissertation presents each research objective sequentially throughout chapters.

*Introduction and background.* Chapter 1 introduces the research topic and clearly outlines research objectives. Chapter 2 summarizes the current methodologies that quantify CCF effects. This includes a critical examination of limitations and assumptions used in the current methodologies.

*Objective 1: Unified CCF Definition.* Chapter 3 provides a literature review, analysis and proposal for a unified definition of CCF.

*Objective 2: Failure data taxonomy.* Chapter 4 provides a literature review, analysis of the CCF and proposal for a new failure data taxonomy to support future CCF models.



*Objective 3: Propose a cause-based data-informed CCF model.* Chapter 5 provides a literature review of previously proposed models. Chapter 6 will propose the PAFM, and Chapter 7 will propose the GDM.

*Summary and conclusions.* Chapter 8 summarizes the proposals of this dissertation and documents future work required for implementation and areas of future research.

## Chapter 2: Current CCF Analysis Methodology and its

### Limitations

#### ***2.1. Introduction***

This chapter provides an overview of the current methodology for incorporating CCFs within a PRA. It will focus on key technical areas of the analysis and includes two simple examples. Limitations of the current methodologies will be identified and summarized at the conclusion of this chapter.

There have been many reports detailing procedures and tools required to conduct CCF analysis within a PRA, however the focus of this chapter will be based on current NRC advice (Mosleh et al. 1998; Wierman et al. 2007). This is the most widely referenced CCF analysis methodology and has iteratively matured. (Mosleh 1991; Mosleh et al. 1988; Mosleh et al. 1998; Fleming et al. 1983). The steps involved in conducting a CCF analysis are summarized in Figure 4.

## **1. SCREENING ANALYSIS**

### **1.1 Problem Definition and System Modeling**

- 1.1.1 Plant familiarization
- 1.1.2 Identification of system and analysis boundary conditions
- 1.1.3 Development of component level system fault tree

### **1.2 Preliminary Analysis of CCF Vulnerabilities**

- 1.2.1 Qualitative screening
- 1.2.2 Quantitative Screening

## **2. DETAILED QUALITATIVE ANALYSIS**

### **2.1 Review of Plant Design and Operating Practices**

### **2.2 Review of Operating Experience**

### **2.3 Development of Cause-Defense Matrices**

## **3. DETAILED QUANTITATIVE ANALYSIS**

### **3.1 Common Cause Modeling**

- 3.1.1 Identification of common cause basic events (CCBEs)
- 3.1.2 Incorporation of CCBEs into fault trees
- 3.1.3 Parametric representation of CCBEs

### **3.2 Data Analysis and Parameter Estimation**

- 3.2.1 Parameter estimation
- 3.2.2 Basic event probability development

### **3.3 System Quantification and Results Interpretation**

- 3.3.1 System unavailability quantification
- 3.3.2 Results evaluation/sensitivity analysis
- 3.3.3 Reporting

**Figure 4: Procedural framework for common cause failure analysis (Mosleh et al. 1998, p.10)**

## 2.2. Describing CCFs

Common Cause Failure result from the existence of two factors:

- Failure Cause
- Coupling Factor

*Failure Cause.* The failure cause is the condition that the component failure can be attributed to. The failure cause affects the frequency of component failure, but by itself does not manifest CCF. The definition of cause can adapted to different levels such as 'proximity cause' and 'root cause'.

- Proximity Cause: The readily identifiable condition that led to component failure. The proximity cause can be regarded as a symptom of the failure cause and not necessarily provide the complete understanding of what led to that failed condition.
- Root Cause. The initiating cause of a causal chain which leads to a proximity cause and the eventual failure of the component.

*Coupling Factor.* The coupling factor is the propagation mechanism that enables failure of multiple components. The coupling factor is a dependency between components which is not explicitly modeled.

*Defenses.* Defenses are the parts of a system that protect against the failure cause or the coupling factor. More robust defenses lower the rate of CCFs.

*Common Cause Component Group (CCCG):* A CCCG is a group of components which share coupling factors, making them susceptible to a common failure cause.

CCFs are only a problem when failures occur within a timeframe that multiple components cannot provide their function. This is sometimes called ‘simultaneous failure.’ It should be noted that the term ‘simultaneous’ is relative. For components that can easily be repaired or replaced, this can be multiple failures within seconds, minutes or hours. For components which can never be replaced or repaired (such as on a space mission) simultaneity can be defined as the mission period which may be years or even decades.

### ***2.3. Problem Definition and System Modeling***

In order to illustrate the CCF analysis procedure, two example systems will be used.

#### ***2.3.1. Example 1: Two Train EDG System***

The first example includes a two train EDG system in which the CCF analysis procedure can be shown in its most simplistic representation. The two generators are in standby to provide emergency power to the same power bus. Only one generator needs to run in order to provide sufficient power. A conceptual diagram is shown in Figure 5. The reliability block diagram for the system is shown in Figure 6. The failure probability for the EDG is assumed to be  $Q_T = 0.006$ . The fault tree with the system

failure rate is shown in Figure 7.

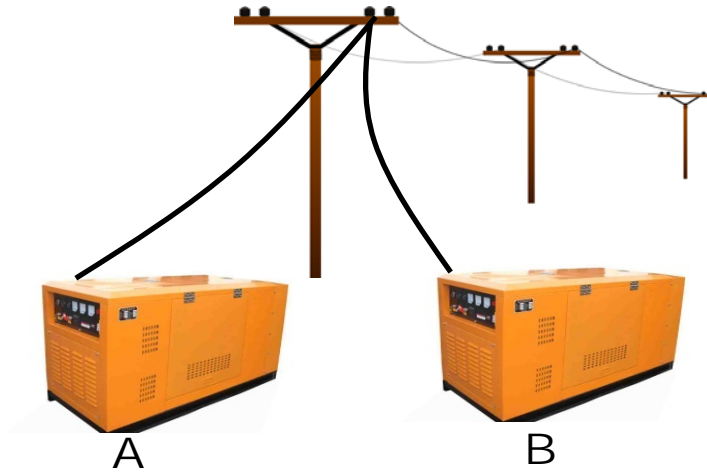


Figure 5: Conceptual diagram for example 1- Two train EDG system

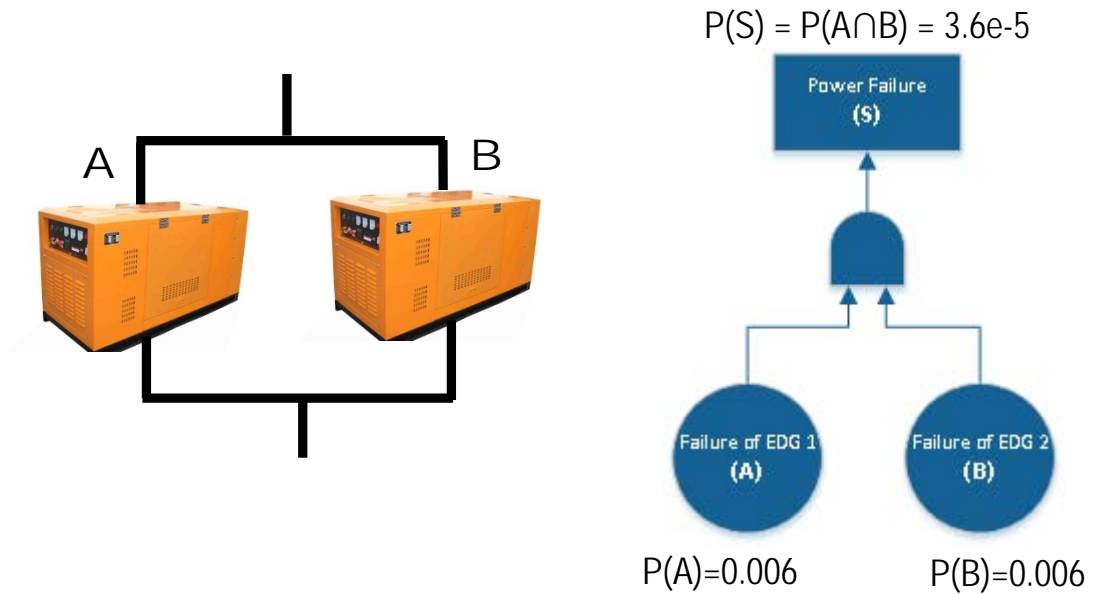


Figure 6: Reliability block diagram for example 1- Two train EDG system

Figure 7: Fault tree for example 1 – Two train EDG system

The cut set for the first example system is:

$$\{A, B\}$$

### ***2.3.2. Example 2: Two Train EDG and Pump System***

The second system consists of a mixture of pumps and generators to highlight the complexity of CCF between component types with varying levels of dependency. The systems objective is to provide water to a cooling system. Only one pump needs to be running in order to provide sufficient water. A pump requires power from an Emergency Diesel Generator to operate. One of the trains has two pumps in redundancy, resulting in a total of three pumps for the system.

The failure probability for an EDG is also assumed to be  $Q_t^{[E]} = 0.006$  and the failure probability for a pump is assumed to be  $Q_t^{[P]} = 0.00204$ . A conceptual diagram of the system is shown in Figure 8. The reliability block diagram is shown in Figure 9. The fault tree with the system failure rate is shown in Figure 10.

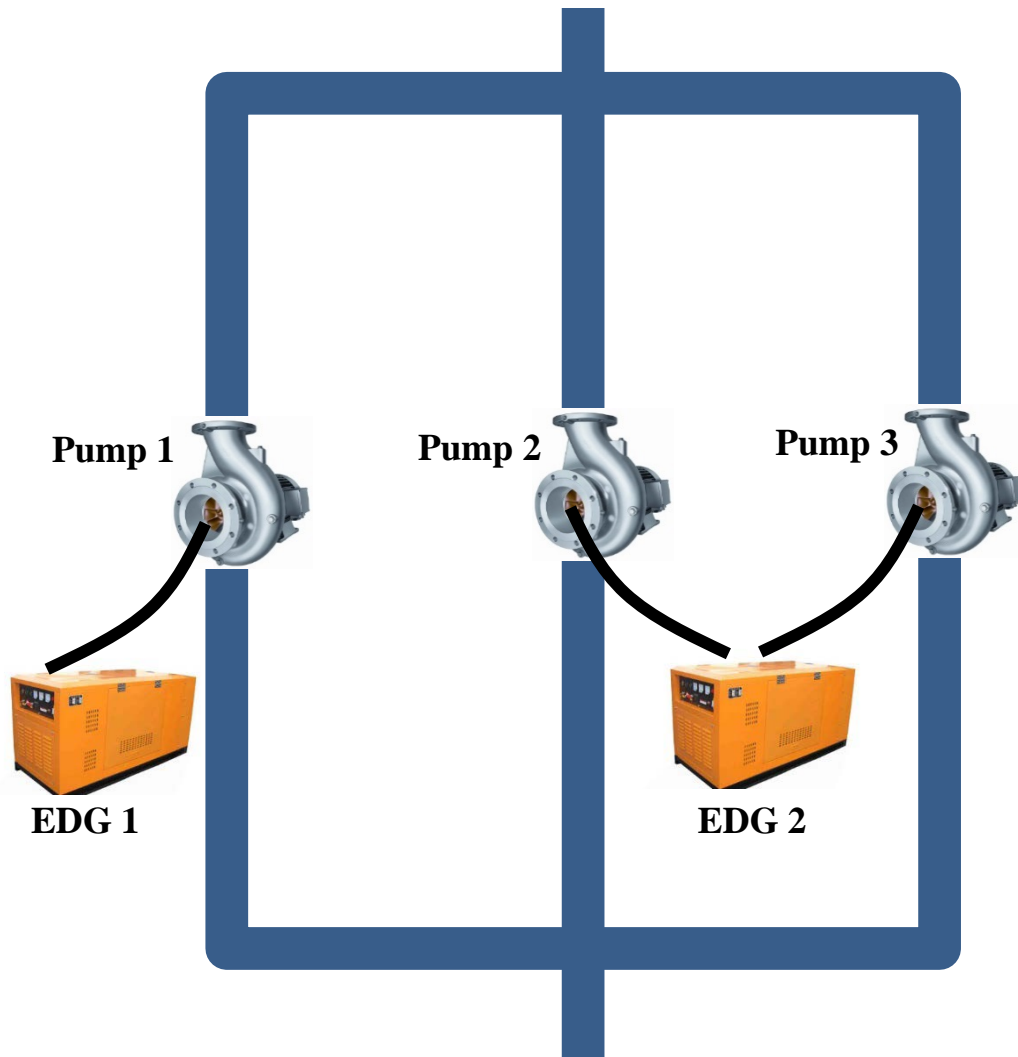


Figure 8: Conceptual diagram for example 2- Two EDGs and three pump system



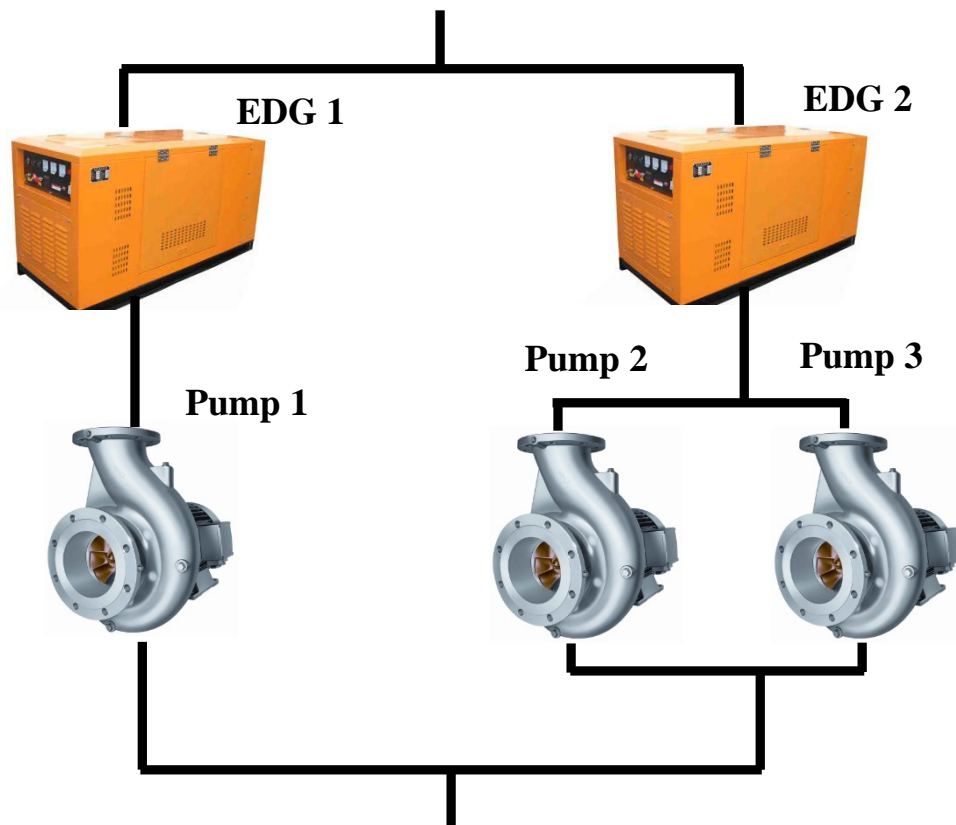
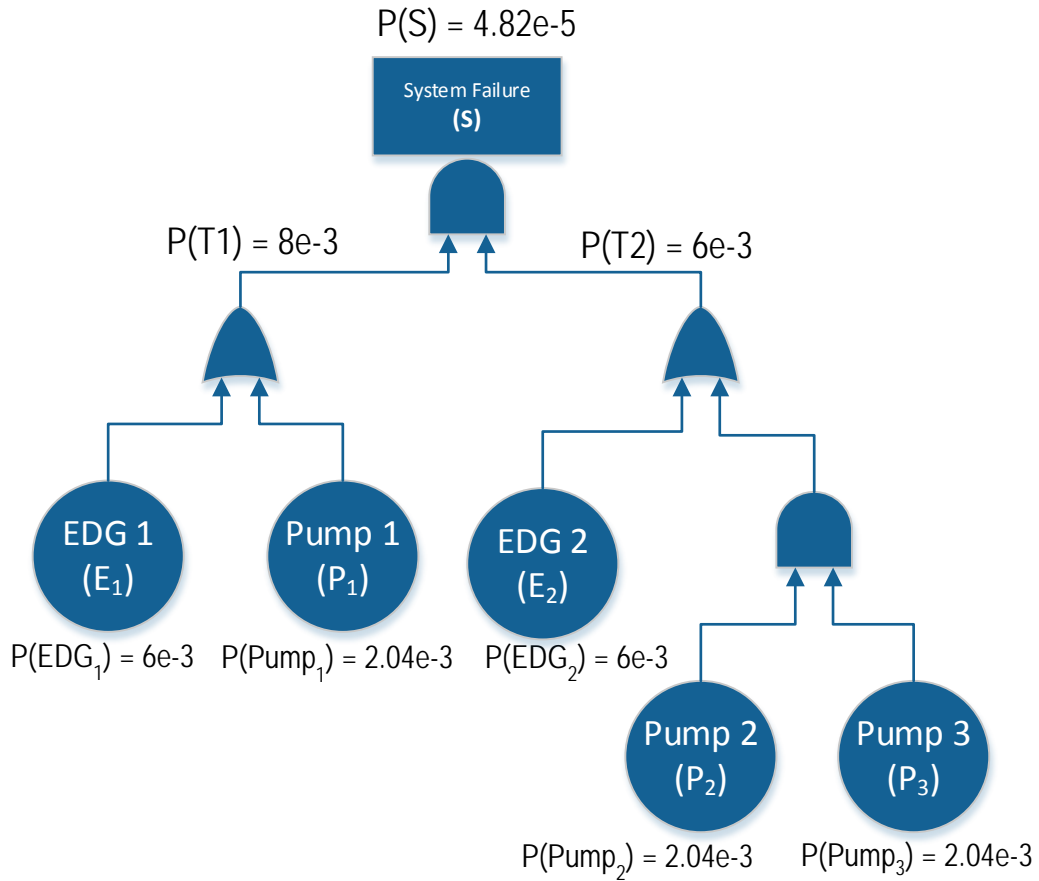


Figure 9: Reliability block diagram for example 2- Two EDGs and three pump system



**Figure 10: Fault tree for example 2 – Two train EDG and pump system**

The cut sets for the second example system are:

$$\{E_1, E_2\}; \{P_1, E_2\}; \{P_1, P_2, P_3\}; \{E_1, P_2, P_3\}$$

## 2.4. Preliminary Analysis of CCF Vulnerabilities

### 2.4.1. Qualitative Screening

Qualitative screening identifies potential vulnerabilities of the system and its components to CCF. This includes identifying the coupling factors that create a soft

dependency between components. The specific coupling factors which should be identified will be discussed in Chapter 3, however NUREG/CR-5485 (Mosleh et al. 1998) recommends identifying components that share one or more of the following:

- design,
- hardware,
- function,
- installation, maintenance or operations staff
- procedures,
- interface,
- location, or
- environment.

Any group of components which share similarities in one or more of these characteristics are assessed for CCF susceptibility. A group of components identified in this process is called a Common Cause Component Group (CCCG).

For simplicity, these example systems will be assessed only for similar features in their install procedures, maintenance staff and location. Example 1 and example 2's assessment is given in

Table 3 and Table 4 respectively.

**Table 3: Qualitative dependency assessment for example 1**

<b>Component</b>	<b>Install Procedure</b>	<b>Maintenance Staff</b>	<b>Location</b>
EDG 1 (A)	EDG IP	Team X	Room Y
EDG 2 (B)	EDG IP	Team X	Room Y

**Table 4: Qualitative dependency assessment for example 2**

<b>Component</b>	<b>Install Procedure</b>	<b>Maintenance Staff</b>	<b>Location</b>
EDG 1 (E1)	EDG	Team X	Room Y
EDG 2 (E2)	EDG	Team X	Room Y
Pump 1 (P1)	Pump V1.1	Team X	Room Y
Pump 2 (P2)	Pump V2.8	Team X	Room Y
Pump 3 (P3)	Pump V1.1	Team Y	Room X

At this point, the objective is a binary decision that concludes whether a set of components should form a CCCG based on their dependencies. Although not stated explicitly in this step, it will become evident that due to the assumption that a CCCG shares all coupling factors, it is impossible for a component to be a member of multiple CCCGs.

For example 1, it is obvious that EDG 1 and EDG 2 share all coupling factors and so it is appropriate that they form a CCCG.

$$CCCG^{[E]} = \{A, B\}$$

For example 2, it is also obvious that EDG 1 and EDG 2 share all coupling factors and should form a CCCG. However, pump 1 and pump 2 share all coupling factors less the installation procedure. Pump 3 only has the installation procedure in common with pump 1. Furthermore, the EDGs and pumps are in the same location and exposed to the same environment, and also exposed to the same maintenance team. It is clear that components within this system share some dependencies, but not others. It is not clear

how the components should be divided into discrete CCCGs.

Using the current methodology the EDGs would form one CCCG with pumps 1 and 2 forming another despite the cross dependencies between components and the difference between the pumps.

$$CCCG^{[E]} = \{E_1, E_2\}$$

$$CCCG^{[P]} = \{P_1, P_2\}$$

**Limitation 1: The current methodology does not account for partial dependencies between components.**

**Limitation 2: Due to limitation 1, the current methodology does not allow a component to be a member of multiple CCCGs.**

#### ***2.4.2. Quantitative Screening***

During quantitative screening, a simple CCF model is applied to the CCCGs identified in the previous step to determine which groups have no significant contribution to the system failure probability.

These are usually components which are in series, or components which already have failure probabilities that are orders of magnitude less than the system failure probability.

For the examples, it is assumed that the CCCGs are significant elements of the system failure probability.

## **2.5. Detailed Qualitative Analysis**

The next step is to conduct a detailed review of the plant specific context in order to tailor the CCF model to the specific system. This involves a qualitative assessment which uses the same basic methodology as the preliminary step, but involves more detail. A documented procedure and inspection checklist to records the potential failure causes, coupling mechanism and defenses specific to the system of interest.

For the example systems, no further detail is required.

## **2.6. Detailed Qualitative Analysis**

### ***2.6.1. Identification of Common Cause Basic Events (CCBEs)***

A CCBE is an event involving failure of a specific set of components due to a common cause (Mosleh et al. 1998, p.41). For example, in  $CCCG^{[E]}$  the basic event A is made up of the contributions from singular/independent failure events,  $A_i$ , and common cause failure events,  $X_{AB}$ , which involve both components A and B.  $A_i$  and  $X_{AB}$  are the CCBEs.

The CCBE events for example 1 and two are shown in Table 5 and Table 6 respectively.

**Table 5: CCBE for example 1**

<b>Component</b>	<b>Common Cause Basic Events</b>
EDG 1 (A)	$A_i, X_{AB}$
EDG 2 (B)	$B_i, X_{AB}$

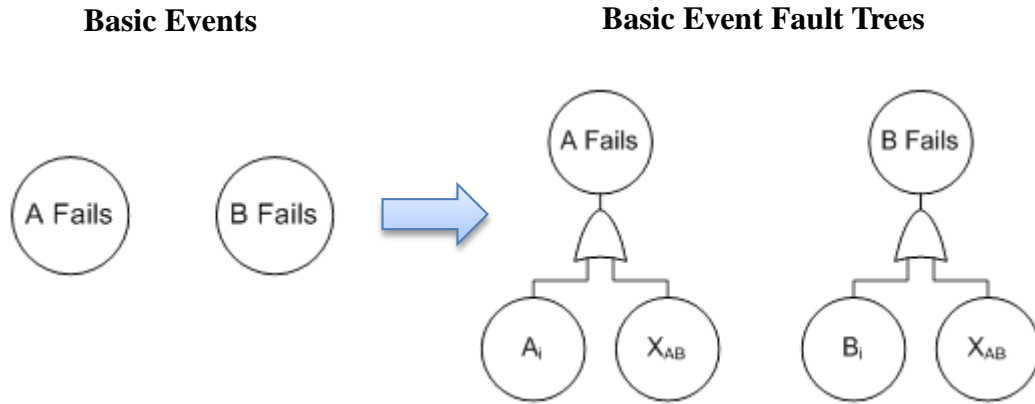
**Table 6: CCBE for example 2**

<b>Component</b>	<b>Common Cause Basic Events</b>
EDG 1 ( $E_1$ )	$E_{1,i}, X_{E_1,E_2}$
EDG 2 ( $E_2$ )	$E_{2,i}, X_{E_1,E_2}$
Pump 1 ( $P_1$ )	$P_{1,i}, X_{P_1,P_2}$
Pump 2 ( $P_2$ )	$P_{2,i}, X_{P_1,P_2}$
Pump 3 ( $P_3$ )	$P_3$

In example 2, limitations 1 and 2 result in CCBEs that does not recognize simultaneous failure of an EDG and pump due to an extreme external environment or maintenance human error. Furthermore the coupling factors between pump 1 and 3 have been ignored.

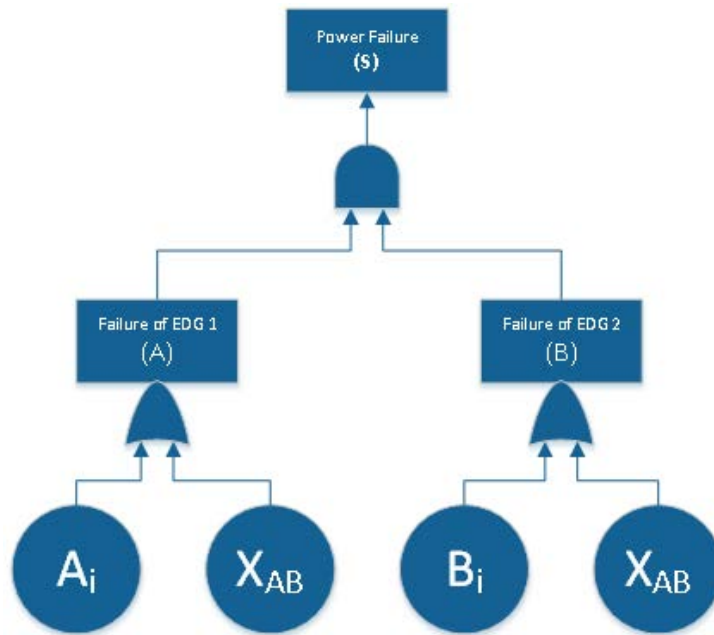
### **2.6.2. Incorporate CCBEs into fault trees**

In this step, the basic events are substituted with the CCBEs. For example the component level fault trees for events A and B (example 1) are shown in Figure 11.



**Figure 11: Conversion of Basic Events to CCBE**

The fault tree for example 1 after substitution of CCBEs is shown in Figure 12.



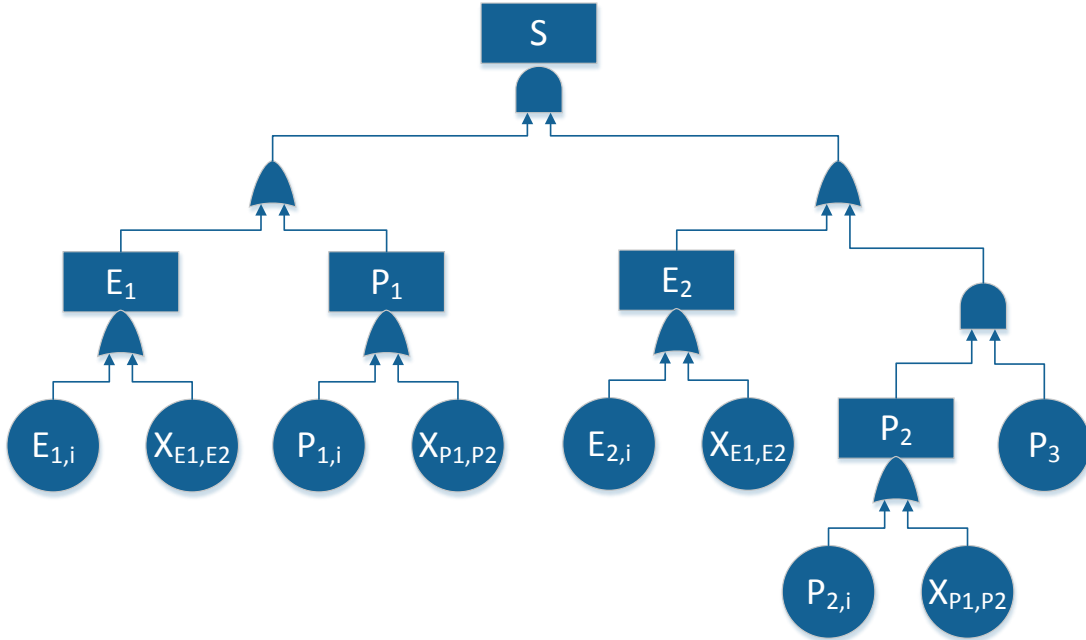
**Figure 12: Fault tree for example 1 with CCBEs**

The cut sets for example 1 are now:

$$\{A_i, B_i\}; \{X_{AB}\}$$

The fault tree for example 2 after substitution of CCBEs is shown in Figure 13.





**Figure 13: Fault tree for example 2 with CCBEs**

The cut sets for example 2 are now:

$$\{E_{1,i}, E_{2,i}\}; \{P_{1,i}, E_{2,i}\}; \{P_{1,i}, P_{2,i}, P_3\}; \{E_{1,i}, P_{2,i}, P_3\}; \{P_3, X_{P1,P2}\}; \{E_{2,i}, X_{P1,P2}\}; \{X_{E1,E2}\}$$

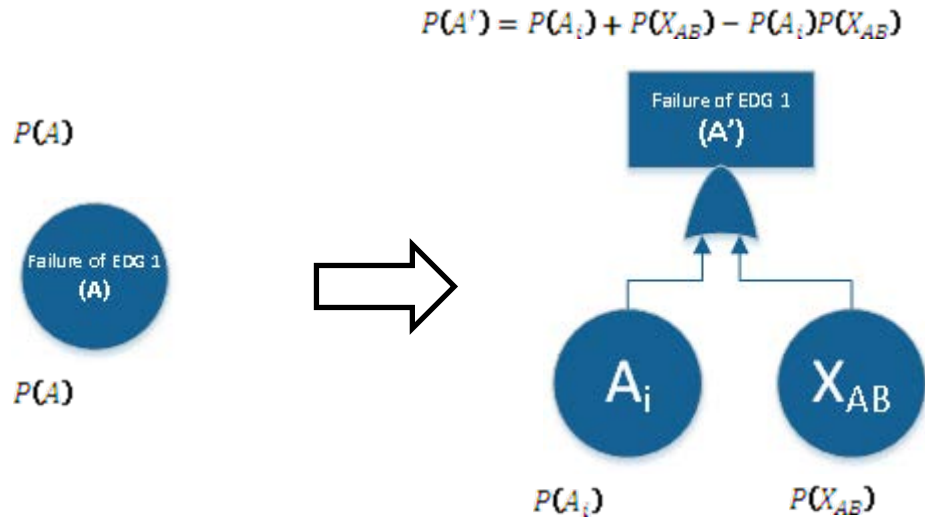
For larger fault trees, the inclusion of the CCBEs can lead to a proliferation of cut sets.

This encourages the analyst to rely on software to solve the augmented fault tree.

**Limitation 3: CCBEs are modeled as independent events instead of mutually exclusive events.**

When events are modeled through the use of an OR gate, they are assumed to be independent events. However the CCBEs are mutually exclusive (i.e. not independent), therefore this modeling technique relies on rare event approximation for accuracy.

For example, the basic event fault tree for event A is shown in Figure 14.



**Figure 14: Basic Event Fault Tree for A**

The assumption driving the identification of CCBEs is that the event A results from its singular/independent failure events or its common cause failures events. However we can see that the fault tree representation, A', includes an additional term to adjust for independent, non-mutually exclusive event:

$$P(A') = P(A_i) + P(X_{AB}) - P(A_i)P(X_{AB}) \neq P(A)$$

Where the terms,  $P(A_i)$  and  $P(X_{AB})$  are orders of magnitude smaller than 1, the quantity  $P(A_i)P(X_{AB})$  becomes insignificant. This is a particular instance of rare event approximation.

The approximation can be seen with the following example. Assume:

$$P(A) = Q_t = 0.2, \quad \beta = 0.3$$

$$P(A_i) = (1 - \beta)Q_t = 0.14, \quad P(X_{AB}) = \beta Q_t = 0.06$$

Then the probability of event A can be calculated as:

$$\begin{aligned} P(A) &= P(A_i) + P(X_{AB}) \\ &= 0.14 + 0.06 \\ &= 0.2 \end{aligned}$$

However the basic event fault tree will calculate event A' as:

$$\begin{aligned} P(A') &= P(A_i) + P(X_{AB}) - P(A_i)P(X_{AB}) \\ &= 0.14 + 0.06 - (0.14)(0.06) \\ &= 0.1916 \end{aligned}$$

CCF, by their nature, are very rare events. In almost all cases, the events are rare enough for the difference to be insignificant. Regardless, this is an explicit assumption for this method.

### ***2.6.3. Parametric representation of CCBEs***

The next step involves transformation of system Boolean representation to an algebraic one involving probabilities of the basic events. For example 1, the cut sets can provide a system probability of failure using the following formula:

$$P(S) = P(A_i)P(B_i) + P(X_{AB}) - P(A_i)P(B_i)P(X_{AB})$$

Using rare event approximation, this would be accurately estimated:

$$P(S) = P(A_i)P(B_i) + P(X_{AB})$$

(Mosleh et al. 1998) go on to assume that the probabilities for each component will be

equal through the following formulation:

$$P(A_i) = P(B_i) = Q_1^{(2)}$$

$$P(X_{AB}) = Q_2^{(2)}$$

Where:

$Q_k^{(m)}$  = *basic event failure frequency/probability for k components failing within a common cause component group of size m, ( $1 \leq k \leq m$ ).*

$Q_k^{(m)}$  is a parameter to the Basic Parameter (BP) CCF model which will be discussed further in Chapter 4. Using the assumption of symmetry and rare event approximation, the system failure probability for the example 1 system can be written as:

$$P(S) = \left(Q_1^{(2)}\right)^2 + Q_2^{(2)}$$

This step is rarely conducted by hand and instead is evaluated by software. The problem then moves from constructing a system probability equation to quantification of the CCBEs that substitute into the fault tree basic events.

For example 2, the system failure probability using the rare event approximation is:

$$\begin{aligned} P(S) = & P(E_{1,i})P(E_{2,i}) + P(P_{1,i})P(E_{2,i}) + P(P_{1,i})P(P_{2,i})P(P_3) \\ & + P(E_{1,i})P(P_{2,i})P(P_3) + P(P_3)P(X_{P1,P2}) + P(E_{2,i})P(X_{P1,P2}) \\ & + P(X_{E1,E2}) \end{aligned}$$

The probabilities for the CCBE events would be:

$$P(E_{1,i}) = P(E_{2,i}) = Q_1^{(2)[E]}$$

$$P(X_{E1,E2}) = Q_2^{(2)[E]}$$

$$P(P_{1,i}) = P(P_{2,i}) = Q_1^{(2)[P]}$$

$$P(X_{P1,P2}) = Q_2^{(2)[P]}$$

$$P(P_3) = Q_t^{[P]}$$

Substitution into the system equation gives:

$$P(S) = \left(Q_1^{(2)[E]}\right)^2 + \left(Q_1^{(2)[P]} \cdot Q_1^{(2)[E]}\right) \left(1 + Q_t^{[P]}\right) + Q_t^{[P]} \left(Q_1^{(2)[P]}\right)^2 \\ + Q_2^{(2)[P]} \left(Q_t^{[P]} + Q_1^{(2)[E]}\right) + Q_2^{(2)[E]}$$

**Limitation 4: Using the current methodology it is difficult to model asymmetrical components.**

Despite the qualitative analysis showing partial dependencies between components and a detailed assessment of plant specific conditions, the assumption of component symmetry severely restricts the options of including the qualitative findings within the model. Plant specific data may exist showing a different failure probability/rate for two similar components, which becomes difficult to include within this model.

In example 2, the pumps 1 and 2 were not symmetrical as they had been installed using different procedures. This assumption fails to recognize that CCF is less likely to occur due to an incorrect installation procedure between the two pumps. Furthermore, it can

be seen that this formulation continues to ignore the dependency between the EDGs and the pumps stemming from exposure to the same maintenance team and external environment.

(Mosleh et al. 1998) addresses this issue through the addition of extra CCBEs to offset the component symmetry. This approach is valid - however the difficulty comes in the quantification of the parameters where the symmetrical CCBEs need to have certain events removed for quantification. This requires manual modification during parameter estimation and requires the analyst to re-assess the database events for applicability, which can be onerous.

#### ***2.6.4. Alpha Factor Model Parameterization***

The next step is to continue to parameterize the CCBEs using the CCF model of choice. The AFM is the most commonly used model within the US nuclear industry. Accordingly, CCBEs will be parameterized in this dissertation using this method. An overview of the AFM is provided within Chapter 4. Assuming the alpha parameters were estimated from data collected using staggered testing gives<sup>4</sup>:

$$Q_k^{(m)} = \binom{m-1}{k-1}^{-1} \cdot \alpha_k \cdot Q_t$$

---

<sup>4</sup> Assumptions about the data used to estimate the alpha factors can change the basic parameter estimator. This depends on whether data was collected using a non-staggered or staggered testing approach. Different estimators for these occasions are provided in (Mosleh et al. 1998).

where:

$$\binom{m-1}{k-1} = \frac{(m-1)!}{(k-1)!(m-k)!}$$

$Q_k^{(m)}$  = *basic event failure frequency/probability for k components failing within a common cause component group of size m, ( $1 \leq k \leq m$ ).*

$Q_t$  = *total failure frequency/probability of each component due to independent and common cause events.*

$m$  = *the total number of components in the CCG*

$k$  = *the number of components within a CCG for example: k out of m components failure.*

For example 1 the basic events for each component are:

$$Q_1^{(2)} = \alpha_1 Q_t, \quad Q_2^{(2)} = \alpha_2 \cdot Q_t$$

For example 2 the basic events for each component are:

$$Q_1^{(2)[E]} = \alpha_1^{[E]} Q_t^{[E]}, \quad Q_2^{(2)[E]} = \alpha_2^{[E]} \cdot Q_t^{[E]}$$

$$Q_1^{(2)[P]} = \alpha_1^{[P]} Q_t^{[P]}, \quad Q_2^{(2)[P]} = \alpha_2^{[P]} \cdot Q_t^{[P]}$$

## **2.7. Data Analysis and Parameter Estimation**

### ***2.7.1. Parameter Estimation – Impact Vectors***

Now that the structure and parameterization of the model is complete, the next step is to estimate the parameters of the chosen CCF model. In order to do this, data must first be collected and analyzed.

The NRC have established a CCFDB which records CCF events and classifies them

according to an impact vector methodology (explained below). The taxonomy and uncertainty treatment is detailed in (Wierman et al. 2007) and review of the CCFDB and the classification taxonomy will be provided in Chapter 3.

An impact vector is a numerical representation of a CCF event. For a CCCG size of  $m$ , the impact vector has  $m+1$  elements. The impact vector element, denoted by  $F_k$  equals  $1$  if the failure of exactly  $k$  components failed during the event and  $0$  otherwise.

$$I_h = [F_0, F_1, \dots, F_m]$$

Where:

$$I_h = \text{the } h^{\text{th}} \text{ hypotheses for an observed CCF event. Where } 1 \leq h \leq H.$$

For example consider a CCCG size of  $m = 2$ . Possible impact vectors could be: (Wierman et al. 2007, p.57)

- $I_1 = [1, 0, 0]$             No components failed
- $I_2 = [0, 1, 0]$             One and only one component failed
- $I_3 = [0, 0, 1]$             Two components failed due to a shared cause

Where there is uncertainty, the analyst provides a probability for each possible impact vector and aggregates them into a single vector. For example, if two components failed and the analyst was 90% confident it was a CCF (meaning there is 10% confidence that there were two independent failures), the resulting impact vector would be: (Wierman et al. 2007, p.58)



$$\bar{I} = \sum_{h=1}^H w_h I_h$$

$$I_1 = [0, 0, 1], \quad I_2 = [0, 2, 0]$$

$$\bar{I} = 0.9I_1 + 0.1I_2 = [0, 0.2, 0.9]$$

When events involve degraded states, the analyst assesses the degree of degradation,  $p_k$ , that would have led to a failure during a typical mission as defined in the PRA and uses that value instead of 1 in the impact vector. When CCF events are distributed in time between occurrences, the impact vector method adjusts the impact vector values based on a binomial probability model. A detailed description of how impact vectors account for uncertainty is contained in (Wierman et al. 2007, p.58).

Impact vectors can be mapped from generic data to plant specific analysis by assessing the applicability of each vector using a weighting factor. The specific method for impact vector mapping and assumptions in doing so will be discussed in section 0.

The sum of the average impact vectors,  $\bar{I}$ , from  $J$  CCF events can be calculated as:

$$n = [n_0, n_1, \dots, n_m]$$

where

$$n_k = \sum_{j=1}^J \bar{F}_k(j)$$

$n_k$  = *the total number of CCF basic events involving the failure of  $k$  similar component.*

$\overline{F}_k(j) =$  is the  $k^{th}$  element of the average impact vector for the  $j^{th}$  event.

Table 7 contains example data from an EDG. This data set is a simplified version of what is held within the NRC failure databases. The impact vector for each CCF event is shown next to the failure information.

The sum of impact vectors for an EDG is:

$$n^{[E]} = [29400, \quad 343, \quad 7]$$

The total failures are:

$$n_F^{[E]} = \sum_{k=1}^m kn_k = 343 + (2)(7) = 357$$

The total demands are:

$$N_1^{[E]} = m \sum_{k=1}^m n_k = 2(29400 + 343 + 7) = 59500$$

Where:

$N_k =$  The number of demands on a subset group of components within the CCGG of size  $k$ . Assuming each time the system is demanded, all components are demanded gives  $N_k = \binom{m}{k} N_D$ .

The failure rate is:

$$Q_t^{[E]} = \frac{n_F^{[E]}}{N_1^{[E]}} = \frac{357}{59500} = 0.006$$

Table 9 contains example data from a pump. The sum of impact vectors for a pump is:

$$n^{[P]} = [44433, \quad 168, \quad 7]$$

This gives the following quantities:

$$n_F^{[P]} = 182, \quad N_1^{[P]} = 89216, \quad Q_t^{[E]} = 0.00204$$

**Limitation 5: The size of the CCCG for single failures are unknown.**

The population sizes for single failures (shown in red) are not currently recorded in the NRC failure databases. This means the failure may have had no possibility of being a CCF due to a lack of coupling factors, but should it have occurred where identical components were in redundancy, the whole CCCG may have failed. Instead, this value is assumed as the average value from the relevant CCF data points (Wierman & Kvarfordt 2011) which may be conservative or optimistic dependant on the circumstances. This is particularly relevant when mapping data from systems with a different number of components in the CCCG (using impact vector mapping) which will be discussed in section 0.

**Table 7: Example Failure Data for Emergency Diesel Generator<sup>5</sup>**

Failure Data				Impact Vector		
Serial	No. Fail	Pop	Cause	$F_0$	$F_0$	$F_0$
1	2	2	IP	0	0	1
2	1	2	IP	0	1	0
3	1	2	MH	0	1	0
4	1	2	IP	0	1	0
5	1	2	IP	0	1	0
6	1	2	IP	0	1	0
7	1	2	MH	0	1	0
8	1	2	IP	0	1	0
9	1	2	IP	0	1	0
10	1	2	MH	0	1	0
11	1	2	MH	0	1	0
12	1	2	IP	0	1	0
13	1	2	MH	0	1	0
14	1	2	MH	0	1	0
15	1	2	IP	0	1	0
16	2	2	MH	0	1	0.5
17	1	2	MH	0	1	0
...	...	...	...	...	...	...
<i>J</i>	1	2	MH	0	1	0
Demands without failures				29400	0	0
<b>TOTAL</b>				<b>29400</b>	<b>343</b>	<b>7</b>

**Table 8: Summary of Impact Vectors for EDG by Cause**

Cause	$F_0$	$F_0$	$F_0$
IP	0	172.2	2.8
MH	0	154.35	3.15
EE	0	16.45	1.05
No Failure	29400	0	0
<b>Total</b>	<b>29400</b>	<b>343</b>	<b>7</b>

<sup>5</sup> Population sizes for single failure (shown in red) are not currently recorded in the NRC CCF database or RADS database.

**Table 9: Example Failure Data for Pump<sup>6</sup>**

Failure Data				Impact Vector		
Serial	No. Fail	Pop	Cause	$F_0$	$F_0$	$F_0$
1	1	2	IP	0	1	0
2	1	2	MH	0	1	0
3	1	2	EE	0	1	0
4	1	2	EE	0	1	0
5	1	2	EE	0	1	0
6	1	2	MH	0	1	0
7	2	2	MH	0	1.6	0.2
8	1	2	MH	0	1	0
9	1	2	EE	0	1	0
10	1	2	EE	0	1	0
11	1	2	MH	0	1	0
12	1	2	MH	0	1	0
13	1	2	MH	0	1	0
14	1	2	MH	0	1	0
15	1	2	MH	0	1	0
16	2	2	MH	0	0	1
17	1	2	EE	0	1	0
...	...	...	...	...	...	...
<i>J</i>	1	2	EE	0	1	0
Demands without failures				44433	0	0
<b>TOTAL</b>				<b>44433</b>	<b>168</b>	<b>7</b>

**Table 10: Summary of Impact Vectors for Pump by Cause**

Cause	$F_0$	$F_0$	$F_0$
IP	0	26.06625	0.18375
MH	0	59.4125	1.8375
EE	0	82.52125	4.97875
No Failure	44433	0	0
<b>Total</b>	<b>44433</b>	<b>168</b>	<b>7</b>

<sup>6</sup> Population sizes for single failure (shown in red) are not currently recorded in the NRC CCF database or RADS database.

### 2.7.2. Parameter Estimation - Alpha Factor Model

The AFM parameter represents the failure ratios for each multiplicity of failure within the CCCG. Each  $\alpha_k$  factor is the probability that, given a failure has occurred,  $k$  components will fail. The AFM parameters are defined and calculated as (Mosleh et al. 1998, p.76):

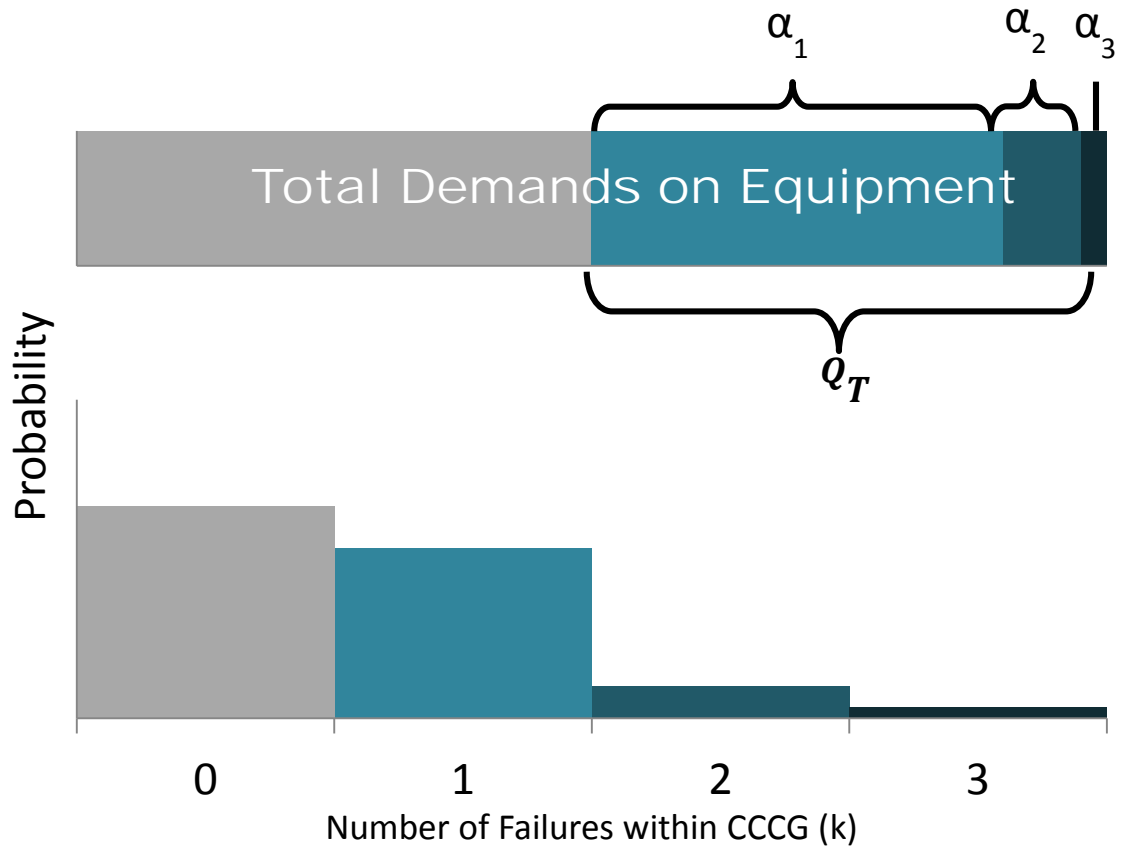
$$\alpha_k = \frac{n_k}{\sum_{j=1}^m n_j}$$

$\alpha_k$  = the fraction of total failure events/frequency that occur in the system resulting in  $k$  out of  $m$  failures.

$m$  = the number of redundant components

$n_k$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

Figure 15 shows the representation of the alpha factors given the system failure data histogram of each failure multiplicity.



**Figure 15: Representation of alpha factor parameters to failure and demand data-set**

For the EDG example the alpha factors can be calculated as:

$$\alpha_1^{[E]} = \frac{343}{343 + 7} = 0.98$$

$$\alpha_2^{[E]} = \frac{7}{343 + 7} = 0.02$$

For the pump example, the alpha factors can be calculated as:

$$\alpha_1^{[P]} = \frac{168}{168 + 7} = 0.96$$

$$\alpha_2^{[P]} = \frac{7}{168 + 7} = 0.04$$

## 2.8. System Quantification and Results Interpretation

### 2.8.1. System unavailability quantification

In this step, parameter estimates are combined with the development of the CCBEs to quantify the system failure/unavailability probability (Mosleh et al. 1998).

For example 1, the system failure probability is:

$$P(S) = \left(Q_1^{(2)}\right)^2 + Q_2^{(2)}$$

Where:

$$Q_1^{(2)} = \alpha_1 Q_t, \quad Q_2^{(2)} = \alpha_2 \cdot Q_t, \quad Q_t = 0.006$$
$$\alpha_1 = 0.98, \quad \alpha_2 = 0.02$$

Substituting these values into the system equation gives a system failure probability of:

$$\begin{aligned} P(S) &= (\alpha_1 Q_t)^2 + \alpha_2 \cdot Q_t \\ &= (0.98 \times 0.006)^2 + (0.02)(0.006) \\ &= 1.668e-4 \end{aligned}$$

Through the inclusion of CCF modeling, the failure probability of the system in example 1 has increased from 3.6e-5 (section 2.3.1) to 1.7e-4. This increase by a factor of 5 demonstrates the significance of CCF.

For example 2, the system failure probability is:



$$P(S) = \left(Q_1^{(2)[E]}\right)^2 + \left(Q_1^{(2)[P]} \cdot Q_1^{(2)[E]}\right) \left(1 + Q_t^{[P]}\right) + Q_t^{[P]} \left(Q_1^{(2)[P]}\right)^2 \\ + Q_2^{(2)[P]} \left(Q_t^{[P]} + Q_1^{(2)[E]}\right) + Q_2^{(2)[E]}$$

Where:

$$Q_1^{(2)[E]} = \alpha_1^{[E]} Q_t^{[E]}, \quad Q_2^{(2)[E]} = \alpha_2^{[E]} \cdot Q_t^{[E]}, \quad Q_t^{[E]} = 0.006$$

$$\alpha_1^{[E]} = 0.98, \quad \alpha_2^{[E]} = 0.02$$

$$Q_1^{(2)[P]} = \alpha_1^{[P]} Q_t^{[P]}, \quad Q_2^{(2)[P]} = \alpha_2^{[P]} \cdot Q_t^{[P]}, \quad Q_t^{[P]} = 0.00204$$

$$\alpha_1^{[P]} = 0.96, \quad \alpha_2^{[P]} = 0.04$$

Substituting these values into the system equation gives a system failure probability of 1.668e-4:

Through the inclusion of CCF modeling, the failure probability of the system in example 2 has increased from 4.82e-5 (section 2.3.2) to 1.668e-4. This is an increase of a factor of 3.5. It should be noted that conservative alpha factors have been used, and increases may be orders of magnitude larger.

### ***2.8.2. Results Evaluation and Sensitivity Analysis***

A sensitivity analysis should now be conducted based around the following areas of uncertainty: (Mosleh et al. 1998)

- system probability on variations of the inputs.
- statistical inference on limited sample size.

- uncertainty on assumptions made in the model.
- uncertainty introduced during data gathering and database development.

This step will not be discussed further, however more detail can be found in (Mosleh et al. 1998).

### ***2.8.3. Reporting***

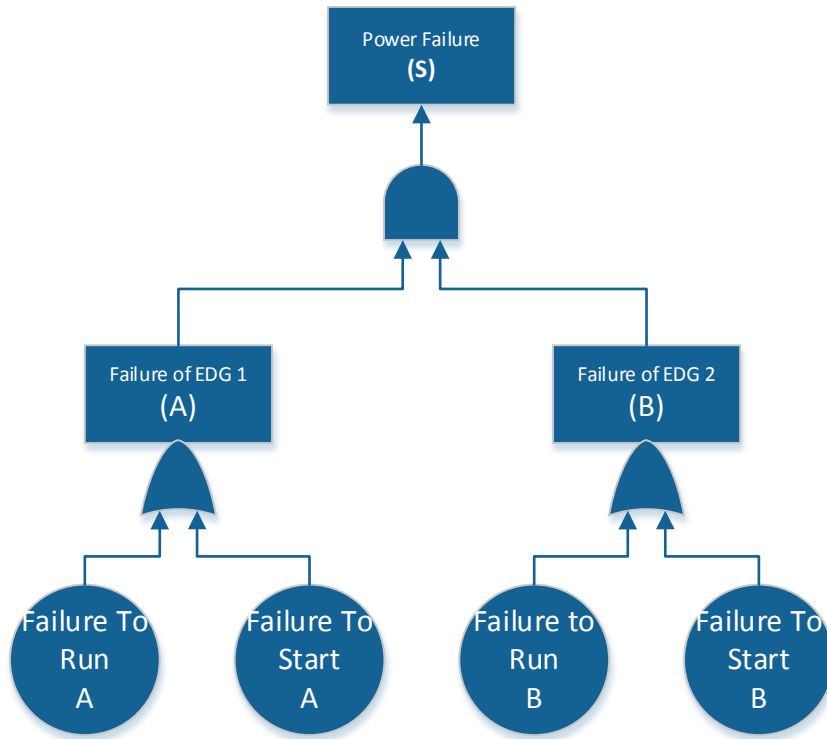
The analysis shall be documented such that its findings can be reproduced, focusing on assumptions and choice of data for parameter estimation. (Mosleh et al. 1998)

### ***2.9. Asymmetrical Components***

It has already been identified that the modeling of asymmetrical components is difficult using the alpha factor method (Limitation 4). This limitation showed that the specific system characteristics assume that all components within the CCCG are identical. The differences in coupling factors such as location, operator and maintenance staff or even manufacturer are not easily accounted for. There is further issue with the assumption of asymmetry between failure modes of the components.

**Limitation 6: Using the current methodology it is difficult to model dependencies between component failure modes.**

The most basic event within a PRA is typically a component failure mode. For example, the two train EDG system fault tree used in example 1 may look like Figure 16.



**Figure 16: Two Train EDG Fault Tree with Failure Modes**

The dependency between the failure modes at the component level is accounted for through the OR Gate. The component will fail regardless of which failure mode occurs. However, the second generator showed strong coupling factors to the first and therefore should be modeled as part of a CCCG. The symmetry restriction requires that each part of the CCCG has the same failure rate. This means the Failure to Run (FTR) and Failure to Start (FTS) modes cannot be within the same CCCG. Instead there will be two CCCGs, a  $CCCG_{FTR}$  and a  $CCCG_{FTS}$ .

The result is an assumption of independence between failure modes. If both generators share the same building and the first generator fails to run due to the room being flooded, the model will believe there is no restriction on the second generator starting.

This problem was described in (Mosleh et al. 1998) when discussing the need to model the asymmetrical situation of service water pumps in standby and continuous operation. This was treated through the addition of an extra basic event to account for asymmetrical features. However, quantification was a task which required reclassification of CCF events. (Jo 2005) suggested a simplified method that models the standby and running pumps separately, and then added an additional event of failure of all pumps. This method was more suitable to situations where the failure of all pumps was a fraction of the CCF events. (Kang et al. 2009) proposed a method which used primary and secondary components and uses a ratio of symmetrical and asymmetrical events in the quantification.

In each case, the quantification activity relied on manual separation of the dataset into symmetrical and asymmetrical events for classification. The GDM proposed in this dissertation overcomes this issue by using the failure cause classification within the database, therefore allowing a more automated analysis process.

### **2.10. Impact Vector Mapping**

There are two types of differences between the system for which data is available and the target system. These are qualitative and quantitative. Qualitative differences include the assessment of the system features and operating environment, and includes different size CCCGs. (Mosleh et al. 1998). Quantitative differences occur when the system in question contains CCCGs that are larger than those for which data exists, the impact vectors created from the database are required to be ‘mapped’ to the size of the target system. Mapping is a complex task, therefore this section will only summarize the ‘mapping up’ process. For more information refer to (Mosleh et al. 1998; Vaurio 2007)

‘Mapping up’ is required when the target system is using impact vectors from smaller sized CCCGs. The example used will be to map from a 2 component CCCG ( $m = 2$ ) to a 4 component CCCG ( $j = 4$ ), where  $m$  is the size of data vector, and  $j$  is the size of the target system). So the question is, given a set of data from a 2 component group, what would the data look like if it had been a 4 component group.

$$\left[ F_0^{(2)}, F_1^{(2)}, F_2^{(2)} \right] \rightarrow \left[ F_0^{(4)}, F_1^{(4)}, F_2^{(4)}, F_3^{(4)}, F_4^{(4)} \right]$$

#### ***2.10.1. Mapping Up Independent Events***

The mapping up of independent events is straightforward. If there are  $n_1^{(2)}$  independent events in a two component system, then this should double if there were twice the number of components exposed (four components)-

$$n_1^{(j)} = \frac{j}{m} n_1^{(m)}$$

$$n_1^{(4)} = \frac{4}{2} n_1^{(2)}$$

### ***2.10.2. Mapping Up :Lethal Shocks***

The NRC failure event database includes a classification that determines if an event was a lethal or non-lethal shock. For a lethal shock, it is assumed that the whole CCCG would fail, regardless of size. Therefore:

$$n_j^{(j)} = n_m^{(m)} \text{ for lethal shocks}$$

$$n_2^{(2)} = n_4^{(4)} \text{ for lethal shocks}$$

### ***2.10.3. Mapping Up Non-Lethal Shocks***

In order to map non-lethal shocks, a binomial distribution is used where  $\rho$  is the conditional probability of each component failure given a shock. When mapping up, it is assumed that the parameter  $\rho$  remains unchanged. For brevity, only the calculation 'to'  $n_2^{(4)}$  will be calculated. This is done using the following procedure: (Mosleh et al. 1998)

The general form of the Binomial Failure Rate (BFR) equation is:

$$n_k^{(m)} = \mu(1 - \rho)^{m-k} \rho^k$$

Where:

$m =$  is the size of the of the CCGG

$k =$  is the number of failures

The BFR equations for 'from' vector are (Mosleh et al. 1998, pp.C-13):

$$n_0^{(2)} = \mu(1 - \rho)^2$$

$$n_1^{(2)} = 2\mu(1 - \rho)\rho$$

$$n_2^{(2)} = \mu\rho^2$$

The BRF equations for the 'to' vector are (Mosleh et al. 1998, pp.C-13):

$$n_0^{(4)} = \mu(1 - \rho)^4$$

$$n_1^{(4)} = 4\mu(1 - \rho)^3\rho$$

$$n_2^{(4)} = 6\mu(1 - \rho)^2\rho^2$$

$$n_3^{(4)} = 4\mu(1 - \rho)\rho^3$$

$$n_4^{(4)} = \mu\rho^4$$

The  $n_2^{(4)}$  equations can be reformulated as contributions from the  $n_i^{(2)}$  terms where  $0 \leq i \leq m$ . Table 11 compares the basic events from a 4 train system and a 2 train system to determine how often the 4 train elements would have been affected the 2 train elements.

**Table 11: Comparison of 4 train and 2 train system basic events**

<b>Basic Events in 4 Train Sys</b>	<b>Basic Events in 2 Train Sys</b>
A	A
B	B
C	None
D	None
AB	AB
AC	A
AD	A
BC	B
BD	B
CD	None
ABC	AB
ABD	AB
ACD	A
BCD	B
ABCD	AB

From the table we can determine that  $n_2^{(4)}$  is made up of  $\frac{4}{5}$  from  $n_1^{(2)}$  and  $\frac{1}{5}$  from  $n_2^{(2)}$ .

This is because if the data from a 2 train system is used, in a four train system the following events would have shown up as two failures AB, A, A, B, B, and None.

**Limitation 7: In impact mapping up, the contribution from demands has not been included as possible failure events.**

It should be noted, that the mapping from  $n_0^{(2)}$  has not been included. In a two train system, some events may not have triggered any failure, that in a four train system may have failed one or more components. The relevant information is available, and so this issue can be rectified through a procedural change.



Recall that:

$$n_2^{(4)} = 6\mu(1 - \rho)^2 \rho^2$$

This equation can be reorganized such that it may be a function of the  $n_i^{(2)}$  events by splitting into the portions 4/5 as observed from  $n_1^{(2)}$  and 1/5 as observed from  $n_2^{(2)}$ .

$$n_2^{(4)} = \underbrace{\frac{4}{5}n_2^{(4)}}_{\text{function of } n_1^{(2)}} + \underbrace{\frac{1}{5}n_2^{(4)}}_{\text{function of } n_2^{(2)}}$$

Substituting  $n_2^{(4)} = 6\mu(1 - \rho)^2 \rho^2$  into this equation gives:

$$n_2^{(4)} = \underbrace{\frac{24}{5}\mu(1 - \rho)^2 \rho^2}_{\text{function of } n_1^{(2)}} + \underbrace{\frac{6}{5}\mu(1 - \rho)^2 \rho^2}_{\text{function of } n_2^{(2)}}$$

Now to make the first term as a function of  $n_1^{(2)}$  and second term a function of  $n_2^{(2)}$ .

$$n_2^{(4)} = \frac{\frac{24}{5}\mu(1 - \rho)^2 \rho^2}{n_1^{(2)}} n_1^{(2)} + \frac{\frac{6}{5}\mu(1 - \rho)^2 \rho^2}{n_2^{(2)}} n_2^{(2)}$$

$$n_2^{(4)} = \frac{\frac{24}{5}\mu(1 - \rho)^2 \rho^2}{2\mu(1 - \rho)\rho} n_1^{(2)} + \frac{\frac{6}{5}\mu(1 - \rho)^2 \rho^2}{\mu\rho^2} n_2^{(2)}$$

$$n_2^{(4)} = \frac{12}{5}(1 - \rho)\rho n_1^{(2)} + \frac{6}{5}(1 - \rho)^2 n_2^{(2)}$$

The final result for the mapping rule is<sup>7</sup>:

$$n_2^{(4)} = \frac{12}{5}(1 - \rho)\rho n_1^{(2)} + \frac{6}{5}(1 - \rho)^2 n_2^{(2)}$$

## **2.11. Event Assessment**

### ***2.11.1. Event Assessment Using AFM***

Event assessment is an application of PRA in which observed equipment failures and outages are mapped into the risk model to obtain a numerical estimate of the event's risk significance (Kelly et al. 2011). In conducting the event assessment, the PRA model has the observed failures instantiated as failed and the observed successes are left as possibilities.

For every mission time the possible outcomes from the CCCG range from no failures to all failed. After a failure occurs, the outcome of having zero failures ceases to be a possibility. Therefore the remaining options must have their probabilities normalized, according to the following rule, as shown in Figure 17 and Figure 18.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

---

<sup>7</sup> Note that this result is different to the mapping rule provided in (Mosleh et al. 1998; Mosleh et al. 1988)  $P_2^{(4)} = \frac{5}{2}\rho(1 - \rho)P_1^{(2)} + (1 - \rho)^2 P_2^{(2)}$ .

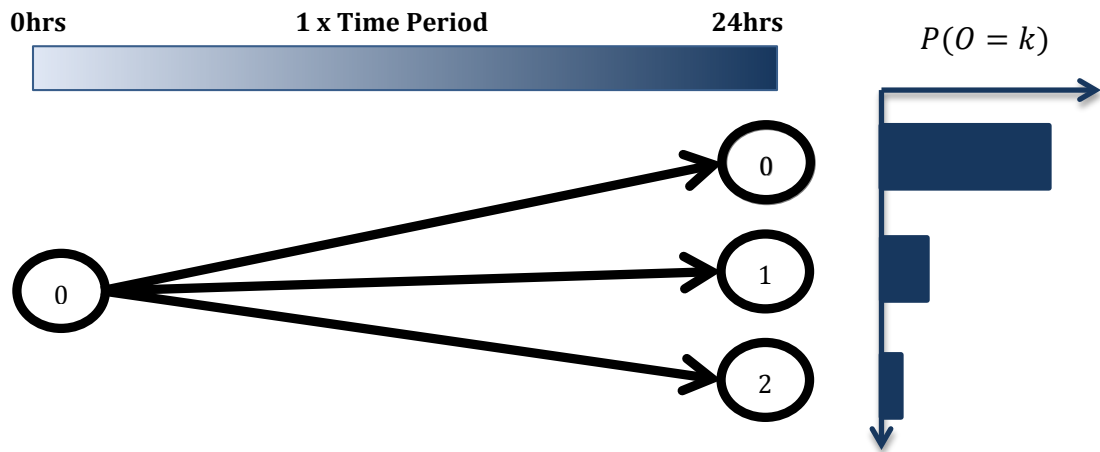


Figure 17: : The probability distribution for the number of failures at the end of a mission (O).

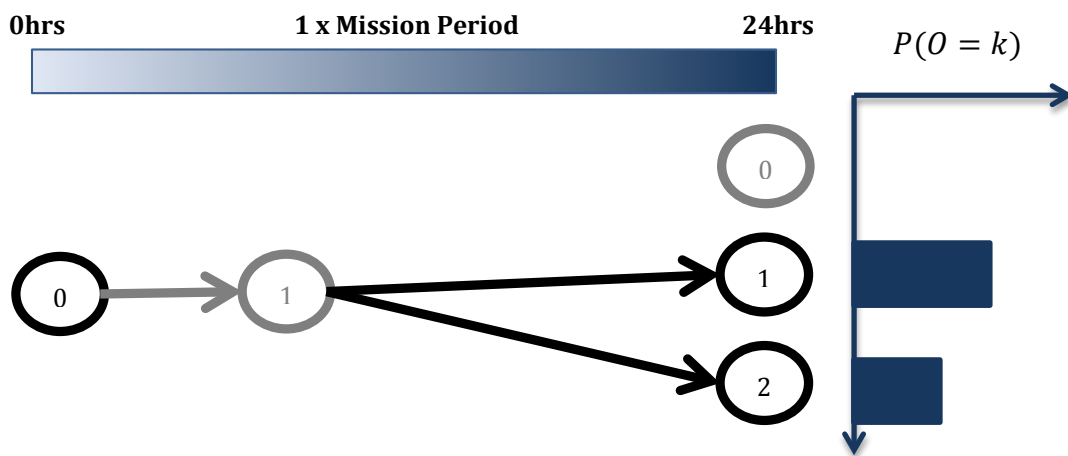


Figure 18: The probability distribution for the number of failures at the end of a mission time (O), after a failure has been observed.

The system failure probability for example 1 is given as:

$$P(S) = \alpha_1^2 Q_T^2 + \alpha_2 Q_T$$

If we assume that component B fails, then the conditional probability for S given B is:

$$P(S|B) = \frac{P(A \cap B)}{P(B)}$$

B can fail from two of the cut sets provided:

$$P(A_i \cap B_i|B) = \frac{P(A_i \cap B_i)}{P(B)} = \frac{Q_1^{(2)}Q_1^{(2)}}{Q_T}$$

$$P(X_{AB}|B) = \frac{P(X_{AB})}{P(B)} = \frac{Q_2^{(2)}}{Q_T}$$

Summing these together we get:

$$\begin{aligned} P(S|B) &= P(A_i \cap B_i|B) + P(X_{AB}|B) \\ &= \frac{Q_1^{(2)}Q_1^{(2)} + Q_2^{(2)}}{Q_T} \\ &= (\alpha_1)^2 Q_T + \alpha_2 \end{aligned}$$

Substituting in parameter values gives:

$$\begin{aligned} P(S|B) &= (\alpha_1)^2 Q_T + \alpha_2 \\ &= (0.98)^2 (0.006) + 0.02 \\ &= 0.02576 \end{aligned}$$

The probability of system failure for example 1 has increased from 1.546e-4 to 0.02576 with knowledge that component B has failed. If components A and B were independent, the probability of system failure would have been 0.006 which would have significantly under-estimated the event assessment.

For example 2, if component  $P_1$  fails, then the conditional probability for S given  $P_1$  is:

$$P(S|P_1) = \frac{P(S \cap P_1)}{P(P_1)}$$

The calculation for each cutset is shown in Table 12.

**Table 12: Cut Sets for Example 2 in event assessment**

Cut Set	$\frac{P(S \cap P_1)}{P(P_1)}$	Boolean Reduction	Basic Parameter
$\{E_{1,i}, E_{2,i}\}$	$\frac{P(E_{1,i} \cap E_{2,i} \cap P_1)}{P(P_1)}$	$P(E_{1,i} \cap E_{2,i})$	$(\alpha_1^{[E]} Q_t^{[E]})^2$
$\{P_{1,i}, E_{2,i}\}$	$\frac{P(P_{1,i} \cap E_{2,i} \cap P_1)}{P(P_1)}$	$\frac{P(P_{1,i} \cap E_{2,i})}{P(P_1)}$	$\alpha_1^{[P]} \alpha_1^{[E]} Q_t^{[E]}$
$\{P_{1,i}, P_{2,i}, P_3\}$	$\frac{P(P_{1,i} \cap P_{2,i} \cap P_3 \cap P_1)}{P(P_1)}$	$\frac{P(P_{1,i} \cap P_{2,i} \cap P_3)}{P(P_1)}$	$(\alpha_1^{[P]} Q_t^{[P]})^2$
$\{E_{1,i}, P_{2,i}, P_3\}$	$\frac{P(E_{1,i} \cap P_{2,i} \cap P_3 \cap P_1)}{P(P_1)}$	$P(E_{1,i} \cap P_{2,i} \cap P_3)$	$\alpha_1^{[E]} Q_t^{[E]} \alpha_1^{[P]} (Q_t^{[P]})^2$
$\{P_3, X_{P1,P2}\}$	$\frac{P(P_3 \cap X_{P1,P2} \cap P_1)}{P(P_1)}$	$\frac{P(P_3 \cap X_{P1,P2})}{P(P_1)}$	$\alpha_2^{[P]} Q_t^{[P]}$
$\{E_{2,i}, X_{P1,P2}\}$	$\frac{P(E_{2,i} \cap X_{P1,P2} \cap P_1)}{P(P_1)}$	$\frac{P(E_{2,i} \cap X_{P1,P2})}{P(P_1)}$	$\alpha_1^{[E]} Q_t^{[E]} \alpha_2^{[P]}$
$\{X_{E1,E2}\}$	$\frac{P(X_{E1,E2} \cap P_1)}{P(P_1)}$	$P(X_{E1,E2})$	$\alpha_2^{[E]} Q_t^{[E]}$

Using rare event approximation and summing the last column of Table 12 gives  $P(S|P_1) = 6.120e-3$  which is an increase from the probability of system failure,  $P(S) = 1.668e-4$ .

### **2.12. Current Issues in CCF Modeling Specific to Event Assessment**

The limitations which have been identified during the CCF analysis are undesirable when conducting event assessments. These limitations are imposed by either the data collection method, the Basic Parameter Model or the Alpha Factor Model.

**Limitation 7: Event assessments do not incorporate knowledge of the cause of failure nor the likelihood of propagation to other components.**

The primary issue is that the conditional probability only accounts for knowledge that a particular component failed. It has no regard for the cause of failure or whether that cause could be propagated through a coupling factor to other components.

### ***2.13. Mission Time***

Probability Risk Assessments (PRA) are made up of many different failure types which are motivated by different life units. For example, a ‘failure to start’ probability is determined by how many start demands are placed on the item. A ‘failure to run’ probability is determined by how many hours the component is required to run. In order to provide consistency across the PRA, a standard mission time is defined.

For the purposes of the NRC scenarios which require the plant to be made safe, the mission period is determined as the time it takes to establish a safe and stable condition. Any failure during that time is undesirable, and the failure of all trains of a redundant system during that time will cause system failure.

The determination of a mission time also affects how CCFs are defined, and subsequently recorded. For example, a PRA defines the mission time as 24 hours, and

a faulty maintenance procedure caused two redundant pumps to fail. The first pump failed today, and the second one failed six days later. It is clear that if this had occurred during a plant shut-down, the dependent failure would not have failed the two train system, as the second pump would have lasted the 24 hours (i.e mission time). However if the mission time is changed to seven days, then the two failures would indeed be classed as a CCF.

A change towards a longer definition of mission time means a transition to higher alpha factors. However, the data which has been recorded within the Common Cause Failure Database, has classified CCF events based on a 24 hour mission time. The impact classification rules allow for uncertainty by incorporating a 'delay' factor. However, as shown in Figure 19, because CCF events have been recorded within the CCF database based on a definition of two or more events occurring within 24 hours, the CCF database cannot assist in the quantification of parameters to PRAs where the mission time is different (for example, up to years for NASA missions).

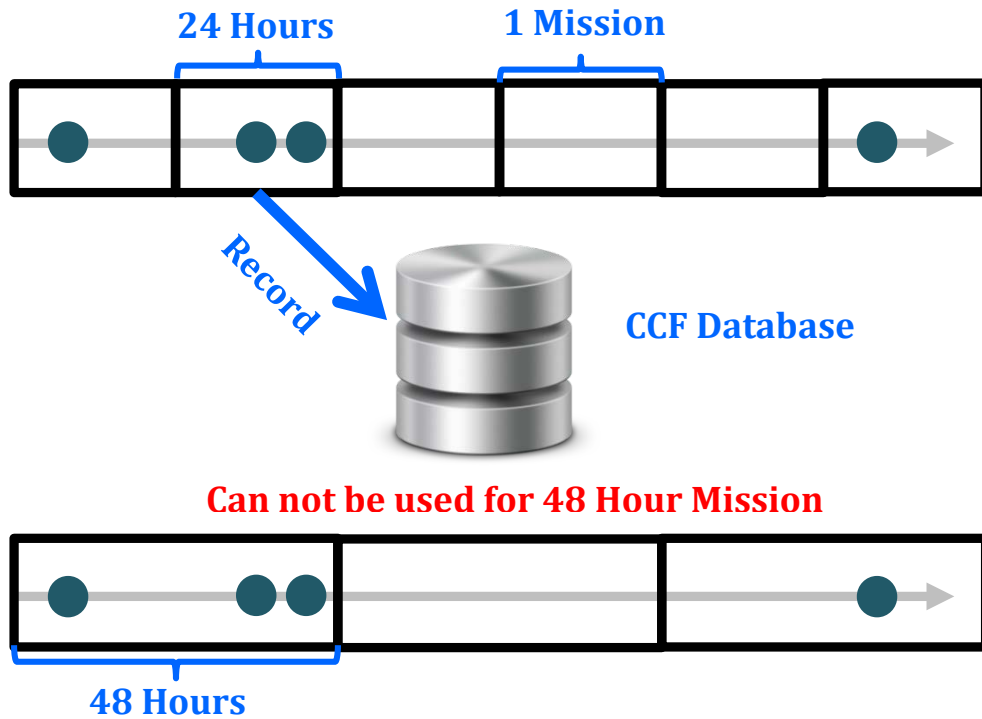


Figure 19: Recording of CCF Events into CCFDB

**Limitation 9: The CCF database is required to have the safe definition of mission time in order to be used for parameter estimation.**

#### 2.14. Summary of Issues

This chapter has provided an overview of the methodology currently used for treating the soft dependencies responsible for the CCF phenomena within a PRA. The methodology shown is based upon the US NRC guidance provided in (Mosleh et al. 1998). Two examples have been used to assist in the explanation of the technique and to help show explicitly the limitations of the techniques. These limitations and the research objectives will form the basis for proposed changes to the CCF analysis methodology.



The primary limitations of the CCF analysis methodology, identified in this chapter, are:

- Limitation 1: The current methodology does not account for partial dependencies between components.
- Limitation 2: Due to limitation 1, the current methodology does not allow a component to be a member of multiple CCCGs.
- Limitation 3: Common Cause Basic Events are modeled as independent events instead of mutually exclusive events.
- Limitation 4: Using the current methodology it is difficult to model asymmetrical components.
- Limitation 5: The size of the CCCG for single failures are unknown.
- Limitation 6: Using the current methodology it is difficult to model dependencies between component failure modes.
- Limitation 7: In impact mapping up, the contribution from demands has not been included as possible failure events.
- Limitation 8: Event assessments do not incorporate knowledge on the cause of failure nor the likelihood of propagation to other components.

- Limitation 9: The CCF database is required to have the same definition of mission time in order to be used for parameter estimation.

## Chapter 3: Definition of Common Cause Failure

### 3.1. Introduction

Despite the proliferation of literature on Common Cause Failure, there are still some inconsistencies and misinterpretation over its definition. This is predominately because the definition of a Common Cause Failure and the scope of Common Cause Failure modeling within a particular system may be different.

The disclaimer at the beginning of NUREG/CR-4780 (Mosleh et al. 1988) demonstrates the lack of consensus regarding the definition of a CCF:

*It is not the purpose of this report to resolve, once and for all, the issues associated with attempts to provide a clear and unambiguous definition of the term common cause event...Here we define what common cause events mean to the system analyst.*

This chapter will cover a literature review of common cause failure definitions, discuss the important characteristics that create the phenomena of common cause failure, and propose a revised CCF definition.

The features of common cause failure which will be discussed are:

- Multiplicity of failure
- Simultaneity of failure
- Functional failures versus component failure
- Independent failure versus single failure
- Explicit modeling of dependency versus implicit modeling
- Redundancy of components
- Symmetry of components

### **3.2. Test Cases**

In order to explore the features of common cause failures, the following test cases have been created.

**Test Case 1.** A mission time is defined as 48 hours. During the first 24 hours period, a pump failed due to an incorrect maintenance procedure cause. During the second 24 hour period the second pump failure occurred also due to an incorrect maintenance procedure. The failure was rectified within minutes. The same procedure was used on the two pumps. Two failures occurred within the mission time, with a shared cause and coupling factor.

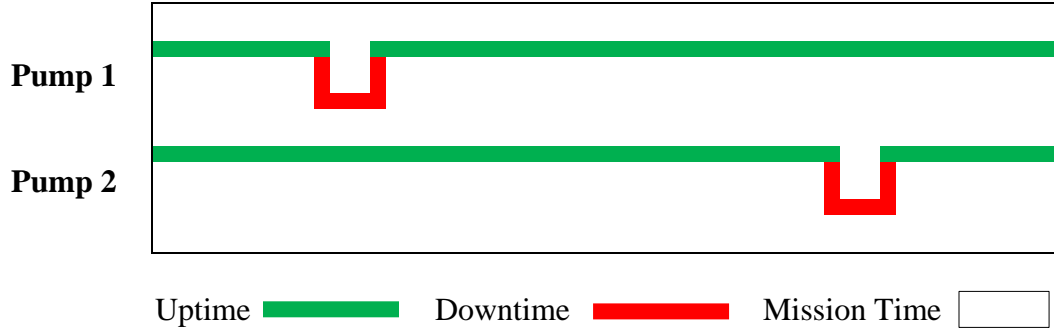


Figure 20: Test Case 1 scenario

**Test Case 2.** A mission time is defined as 24 hours. The same events of test case 1 occur where only one failure occurred during each mission, with a shared cause and coupling factor. The pumps were repaired within 6 hours.

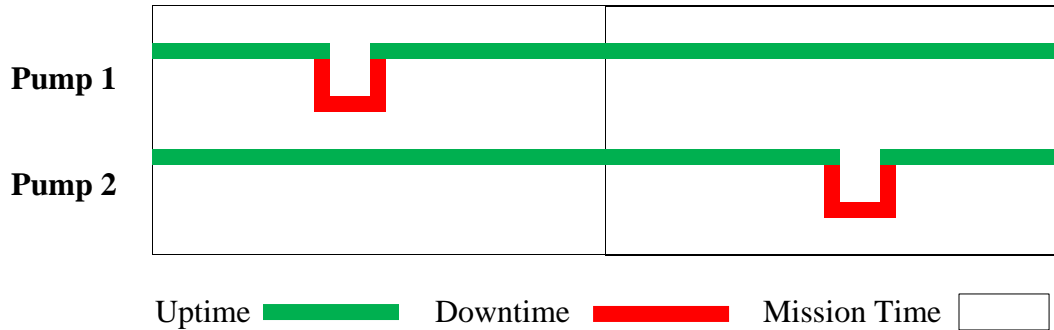
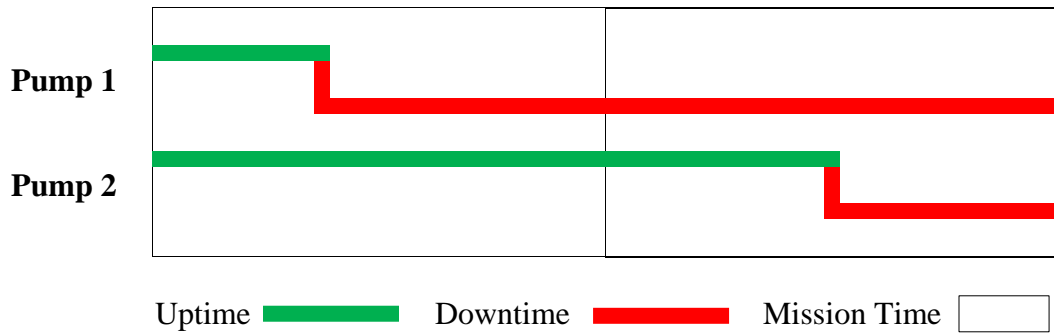


Figure 21: Test Case 2 scenario

**Test Case 3.** A mission time is defined as 24 hours. The same events of test case 1 occur however it takes three days for each pump to be repaired. The two pumps are in redundancy and provide a safety critical function. Therefore during the second mission period, the safety function provided by the pumps failed.



**Figure 22: Test Case 3 scenario**

**Test Case 4.** A mission time is defined as 24 hours. During a mission time two pumps were unable to perform their function, due to loss of AC power. The failure was functional; however the components were operable if AC power could be provided.

**Test Case 5.** A mission time is defined as 24 hours. During a mission a fire occurred which destroyed the building including the two pumps. This failure mechanism and dependency had been explicitly modeled within the PRA.

**Test Case 6.** A mission time is defined as 24 hours. The same events of Test Case 6 occur however the fire and dependency between the pumps had not been explicitly modeled within the PRA.

**Test Case 7.** A mission time is defined as 24 hours. During a mission time, a pump and a generator, which were co-located, failed due to extremely high ambient temperatures within the room.

**Test Case 7.** A mission time is defined as 24 hours. During a mission time, two pumps which were co-located, failed due to extremely high ambient temperatures within the room. One pump provided a safety function to the plant operation, while the second pump cleared storm water. The two pumps were not in redundancy.

**Test Case 8.** A mission time is defined as 24 hours. During a mission time, a pump and a generator, which were co-located, failed due to extremely high ambient temperatures within the room.

### 3.3. Literature Review

In WASH-1400 (Rasmussen 1975) the term Common Mode Failures (CMF) was used as an all-inclusive term. Almost any multiple failure event which are not independent were included in the CMF definition. (Rasmuson et al. 1979) discusses how the term Common Cause Failure is a preferred term and more specific in its meaning.

A detailed literature review has been conducted by Smith and Watson in 1979, collating nine different definitions which had used 12 attributes to describe CCF (Smith & Watson 1980). Each definition was compared against the 12 attributes and found that none of the definitions agreed. Of unanimous agreement between the definitions was:

- the requirement for multiple failures;

- the requirement for a shared cause;
- no requirement that the components actually be challenged during the time it failed;
- no requirement that a component should have the same failure mode (despite the term Common Mode Failure being popular).

In response to this analysis, Smith and Watson proposed a new definition of CCF:

*Inability of multiple, first-line items to perform as required in a defined critical period of time due to a single underlying defect or physical phenomenon such that the end effect is judged to be a loss of one or more systems.*

In an earlier version of NUREG guidance on CCF, NUREG/CR-4790 (Mosleh et al. 1988) the Common Cause Event is defined as:

*In the context of system modeling, common cause events are a subset of dependent events in which two or more component fault states exist at the same time, or in a short time interval, and are a direct result of a shared cause.*

Vaurio suggested that each analyst select the attributes essential for his definition and explain under what titles the other classification of failures have been placed. (Vaurio



1981)

Paula, provided a review of the definition of CCF such that guidance can be provided on the data collection of such events (Paula 1995). Whilst acknowledging that CCF is defined by the analyst and both general and narrow definitions may be true, she added:

*CCF events are dependent failures resulting from causes that are not explicitly modeled.*

The primary reference for this thesis is NUREG/CR-5485 (Mosleh et al. 1998) which defines Common Cause Failure as:

*A CCF event consists of component failures that meet four criteria:*

*(1) two or more individual components fail or are degraded, including failures during demand, in-service testing, or deficiencies that would have resulted in a failure if a demand signal had been received;*

*(2) components fail within a selected period of time such that success of the PRA mission would be uncertain;*

*(3) component failures result from a single shared cause and coupling mechanism; and*

*(4) a component failure occurs within the established component boundary.*

Despite what seems to be a convergence in the definition of CCF, current literature still

uses and misuses these terms, adding to the confusion. For example (Kaufman et al. 2000) uses the term Common Mode Failure and Ericson includes a definition of CMF which is the opposite to its early use (in WASH-1400): (Ericson II 2005)

*“Common Cause Failure (CCF). The failure (or unavailable state) of more than one component due to a shared cause during the system operation.” “A CCF is the simultaneous failure of multiple components due to a common or shared cause.” “CCFs include common mode failures (CMF), but CCF is much larger in scope and coverage.”*

*“Common Mode Failure (CMF). The failure of multiple components in the same mode (NASA 2002). An event, which simultaneously affects a number of elements otherwise, considered to be independent.” “The term CMF, which was used in the early literature and is still used by some practitioners, is more indicative of the most common symptom of the CCF, but it is not a precise term for describing all of the different dependency situations that can result in a CCF event. A CMF is a special case of a CCF, or a subset of a CCF.”*

### 3.4. Definition Discussion

#### 3.4.1. *Multiplicity of Failure*

In order for there to be a common cause failure, there is little dispute that this will involve multiple failures. However, the concept of multiplicity is directly linked to the concept of simultaneity. The difference between a multiple failure in Test Case 1 and a multiple failure in Test Case 2 is the definition of simultaneity.

#### 3.4.2. *Simultaneity*

The presence of a common cause failure can only cause system failure if the components fail in such a way that the functions from multiple components are concurrently unavailable.

Simultaneity has two components to it, the first being the mission time and the second being the rectification time.

#### **Mission Time**

As discussed in section 2.13, the specification of a mission time is critical in order to define simultaneity. For example Test Case 1 and Test Case 2 had exactly the same events occur, however the definition of mission time changed. This means Test Case 1 is a CCF and Test Case 2 is not.

The concern of multiple failures within the mission time may be independent of the rectification time. For example in Test Case 1, the components were repaired within minutes, and therefore it is unlikely the two components would fail simultaneously, however these multiple failures still place an uncertainty on the system to complete the mission, as repair is not certain. Therefore the NUREG/CR-5485 definition includes the condition “success of the PRA mission would be uncertain”.

### **Rectification Time**

The mission time is critical in defining CCF events where components cannot be repaired, or have long repair times and the mission period is short. However, in most circumstances it is not the mission time that provides the definition of simultaneity, it is the rectification time. This makes the simultaneity part of the CCF definition different for each component type and dependent the specific maintenance support system.

For example, Test Case 2 and Test Case 3 had exactly the same events occur except Test Case 3 had a longer repair time. Both test cases are not classified as CCF given the definition in NUREG/CR-5485 (Marshall et al. 1998) however the dependencies between the two pumps caused a system critical failure. Furthermore Test Case 1 had an almost instant repair time, so despite multiple failures occurring during the mission time, it is unlikely this would have affected the provision of the system critical function.

The CCF definition requires the concept that multiple components are unable to perform their function at the same time. This definition is more general than using a mission time and includes the ideas discussed regarding repair times.

#### ***3.4.3. Functional Failure and Physical Failure (System Boundary)***

Components may fail to perform their required functions by either having a failure within its system boundary, or having a failure occur outside its system boundary which fails to provide a necessary condition for the component to work.

The definition provided by NUREG/CR-5485 (Marshall et al. 1998) recognizes this distinction between classes of dependencies, where most simplified definitions are deficient. This is shown through Test Case 4 where the pumps fail to perform their function due to the failure of shared dependency on AC power. This is not a CCF.

The definition of a CCF must have a more than one failure, and therefore the definitions of single failure and independent failure are not required, except for the purposes of modeling.

#### ***3.4.4. Independent and Common Cause Failures***

Many of the models require the separation of independent failure to common cause failures. This cannot be determined from the failure information itself, and is usually assumed to be the single failure observed within a CCF. Is a single failure within a

CCCG an ‘independent failure’ which would have caused only that component to fail, regardless of size of the CCCG, or is it a CCF failure which had the potential to cause other components to fail. Different interpretations of this question have prompted a myriad of different CCF models.

The separation of single and independent failures only becomes a problem when using single failures to estimate the potential for CCF such as during impact vector mapping. If only a single failure occurred, regardless of whether a shared cause existed and there was potential for another component to fail, the fact remains that no common cause failure occurred.

However, when using the information from a single failure to predict CCF behavior, this problem can be solved by acknowledging that failures occur probabilistically in the presence of a cause. A failure cause condition may be present to multiple components, but only one component fails. Each component’s failure rate is the sum of its failures from each cause. Each failure has the possibility of another component failing, should it share the same cause environment.

With this understanding of the problem, the definition of independent failure is a failure, whose failure cause condition cannot be shared by other components. Therefore the definition of an independent failure is a function of the qualitative assessment discussed in section 2.4.1 and the cause of the failure. This definition allows for the

single failure of a component to still be understood as a CCF event.

Using this understanding of independent failures and CCFs, the proposed General Dependency Model does not require a distinction to be made between independent and single failure events for modeling, other than through the qualitative assessment of coupling factors.

By definition a CCF must have a more than one failure, and therefore the definitions of single failure and independent failure are not required, except for the purposes of modeling.

#### ***3.4.5. Explicit and Implicit Dependencies***

The unexpected failure of multiple components occurs because the events leading to the failure had not been anticipated, despite known soft dependencies having existed between the components. As risk modelers learn more about a particular phenomenon (like extreme external environment events) more detailed models are created, and the dependencies between components, relating to that cause, become explicitly included within the model.

This means the definition of Common Cause Failure will change between each PRA analysis, and depends on the level of detail to which the model explicitly includes dependencies between components. This is the first consideration which makes a CCF

definition dependent on the model instead of a definition of the system and interpretation of events. So the question becomes, whether the definition of CCF is a modeling concept, or whether it is a definable event given a system definition.

Test Case 5 and 6 fail both pumps due to a fire.. Test Case 5 had this scenario explicitly modeled, and Test Case 6 did not. In the case of the CCFDB, such events are not recorded within the database, because it is common practice within the nuclear industry for these events to be modeled explicitly. Therefore the data which is recorded within the CCFDB itself is dependent on the modeling standards set by that industry.

In using the definition for data collection and modeling, two options exist:

- *Exclude reference to modeling within the definition.* This is the status quo, which provides a broad definition of CCF which allows for explicit or implicit modeling. This would result in both Test Case 5 and Test Case 6 being classified as CCFs. The activity of modeling external environmental events, or fire events within the nuclear industry would be considered a type of CCF modeling (despite it being explicit treatment of those failure causes). Furthermore the CCFDB only collects a portion of CCF events, those which support implicit CCF modeling.
- *Include reference to modeling within the CCF definition.* By including modeling as part of the CCF definition means that CCF is a modeling concept



used to represent a sub-set of dependent failures. This means an observed event cannot be classified as a CCF unless there is knowledge of the PRA model used to represent the failure. Using this interpretation, Test Case 5 would not be a CCF, and Test Case 6 would be a CCF.

The second option is an important step in reducing confusion over what a CCF is, and will be used in this thesis. It clarifies why in some industries failure due to fires may be classified as CCF and in others it is not. It explains the different sizes of alpha factors for identical systems where the models contain different levels of details. It also makes clear the important assumptions used when calculating generic model parameters to be used in other industries.

The definition of a CCF requires the concept of modeling implicit dependency relationships.

#### ***3.4.6. Redundant Components***

The most common application of CCF modeling is on redundant components. While the concept of CCF directly attacks the advantages of redundancy, it is not exclusive to redundant components. For example, Test Case 7 would be considered a CCF despite the two pumps not being in a redundant configuration.

The definition for CCF does not require the concept of redundancy.

### ***3.4.7. Identical Components***

The most common application of CCF modeling is on identical components. This is because the target of CCF modeling is redundancy and in many cases redundancy is provided through the installation of like components on parallel trains. The second reason is because almost all CCF modeling techniques require the assumption of symmetry in order to simplify the analysis. This is discussed in throughout Chapter 2.

The phenomena of CCF is not restricted to identical components. For example, in Test Case 8 the generators both failed within a mission time due to a shared cause and coupling factor. This can be classified as a CCF event which involved two different component types.

The definition for CCF does not require the concept of component symmetry.

### ***3.5. Proposed CCF Definition***

A CCF event consists of component failures that meet five criteria:

- (1) two or more individual components fail or are degraded, including failures during demand, in-service testing, or deficiencies that would have resulted in a failure if a demand signal had been received;

- (2) components fail within a selected period of time such that multiple components are unable to perform their intended function or success of the PRA mission would be uncertain.
- (3) component failures result from a single shared cause and coupling mechanism; and
- (4) a component failure occurs within the established component boundary.
- (5) the dependency between components has not already been explicitly modeled.

While this definition is very specific, a shortened version is also desirable. The following simple CCF definition is proposed;

*Unexpected simultaneous failure of two or more components due to a shared cause.*

Using this simple definition, the term ‘unexpected’ is to separate explicit and implicit failures, and the word ‘simultaneous’ is left as a descriptive term requiring further definition by the reader.

Table 13 provides a comparison between the classification of the test cases given the NUREG/CR-5485 definition and the proposed definition.

**Table 13: Comparison of CCF definitions for each test case**

<b>Test Case</b>	<b>Test Case Feature</b>	<b>NUREG/CR-5485</b>	<b>Proposed</b>
1	Two failures in mission time with shared cause and coupling factor, with instant repair.	Yes	Yes

2	Two failures, but not within the same mission time.	No	No
3	Two failures, not within the same mission time, but due to extensive repair times cause system unavailability.	No	Yes
4	Functional failure without component failure	No	No
5	Failures explicitly modeled	Yes	No
6	Failure implicitly modeled	Yes	Yes
7	Different components failed	Yes	Yes
8	Components not in redundancy fail	Yes	Yes

## Chapter 4: Common Cause Failure Database Taxonomy

Data for common cause failures is extremely difficult to obtain. Common Cause Failure events are rare events that even after over 20 years of data collection from joint nuclear data collection initiatives, the number of CCF events are in the hundreds (Wierman et al. 2007). Not only are the event infrequent, but it is difficult to detect a CCF even if failure data is abundant. Issues about the scope of CCF, as discussed in chapter 3, and manual review of failure events make data collection resource intensive.

Due to the difficulties in obtaining CCF data, a key feature of all CCF models is to either use qualitative assessments and expert judgment, or maximize the use of the available data through the use of assumptions which may not be possible to verify. Despite quantitative models being proposed during the late 1970s, early 1980s (Fleming 1975; Mankamo 1977; Mosleh & Siu 1987; Vesely 1977) the availability of data to support such models was limited. In 1985 the first generic CCF model parameters were published as EPRI NP-3967, however the study had limitations that prevented the quantification of many models. Over the past 20 years a common goal has been to reach consensus for the methods of recording and classifying data and the collation of data from multiple sources into failure databases.

There is a direct relationship between the way failure data is recorded and the ability to support CCF model quantification. For example in 2004 INL began recording the

failure cause for single failure events, and back-fit these classifications to data from 1997 (Wierman 2013). This is a key information requirement for both the Partial Alpha Factor Model (PAFM) and the General Dependency Model (GDM) proposed within this thesis.

Another key information requirement of the PAFM and GDM is clear relationship between the failure cause and the coupling factor. For example, if a failure occurs and the cause is recorded in the failure database, the Partial Alpha Factor wants to know what coupling factors this failure could have propagated through. The General Dependency Model quantifies the strength of each coupling factor using data from failure causes. This will be a difficult task if failure causes could have propagated through multiple coupling factors and visa versa.

This chapter will assess the NRC Common Cause Failure Database (CCFDB) data to determine if the classification allows for inference of coupling factors from causes and if there is ambiguity between classifications.

#### ***4.1. Literature Review***

The following section will provide a brief literature review of failure taxonomies for common cause failure events. For brevity this review will not compare or list the classification taxonomies, as they have little impact on the objective of this chapter. The classification scheme and definition of terms will change between industries and

needs to be established by industry experts. Particular attention will be paid to the development of the classification scheme currently used within the CCFDB.

In 1979, Hagan summarized the early classification schemes and found general agreement in their approach. In 1979 Edwards and Watson proposed a classification system which divided causes into Design, Construction, Procedural and Environmental, with 46 sub-categories (Edwards & Watson 1979). In 1988 NUREG/CR-4780 summarized the PRA Procedures Guide, EPRI Integration Procedure Guide, the EPRI Event Classification Scheme for their effect on PRA. Many of the classification systems also classified the type of dependency, as the definition for common cause failure was also in a state of flux. All classification schemes had the notion of cause and coupling factor (Mosleh et al. 1988).

Current guidance introduces classification systems for defenses, although these have more variability between taxonomies than the cause and coupling factor classifications. Significant failure classification guidance is currently provided in the following documents:

- NUREG/CR-5485 is the current guidance on modeling common cause failures for the US nuclear industry and contains a coupling factor and failure cause classification system. A matrix is presented which shows the impact of different failure mechanisms against root causes and coupling factors. Many of the failure mechanisms show a one to many relationship. (Mosleh et al. 1998)

- NUREG/CR-6268 Rev 1 provides the current guidance on event data collection, classification and coding for the NRC CCFDB. The classification scheme used in this document will be discussed in more depth during this chapter. (Wierman et al. 2007)
- Unified Partial Method is a popular CCF model within the UK and Europe. This methodology focuses on a qualitative assessment of defenses for causes and coupling factors (Brand & Gabbot 1993). In Zitrou's Bayesian Network modeling of the UPM classification system, a many to many relationship exists between causes and coupling factors and she includes a method to treat dependencies between these factors within her model (Zitrou 2006a). Note that the UPM classification system is used for the UPM model, not for classifying failure data.
- NEA/CSNI/R(2011)12 provides the current guidance for classifying failure data for contribution to the International Common Cause Failure Data Exchange (ICDE). (NEA 2011)

In 2007, Lindberg mapped the current classification approaches together (NRC, ICDE and UPM). She also mapped which causes influence which defenses/coupling factors to create a qualitative tool called, Relations of Defences, Root Causes, and Coupling Factors (RDRC) diagrams. This allowed the combination of data from sources which were classified differently. Furthermore histograms are created from failure data to



determine which defenses or coupling factors would have influenced the most CCF events. This provides qualitative analysis support to decisions on where to invest in defenses against CCF. An example of this diagram is included as Figure 23. The relationship between coupling factors, failure causes and defenses can be seen.

The primary purpose of having a failure and coupling factor classification system has been to provide understanding and insight into why CCF occurs and ways to defend against it. Until the requirements of the two models proposed in this thesis, there has not been a specific requirement to relate the classification concepts except through the conduct of qualitative analysis.

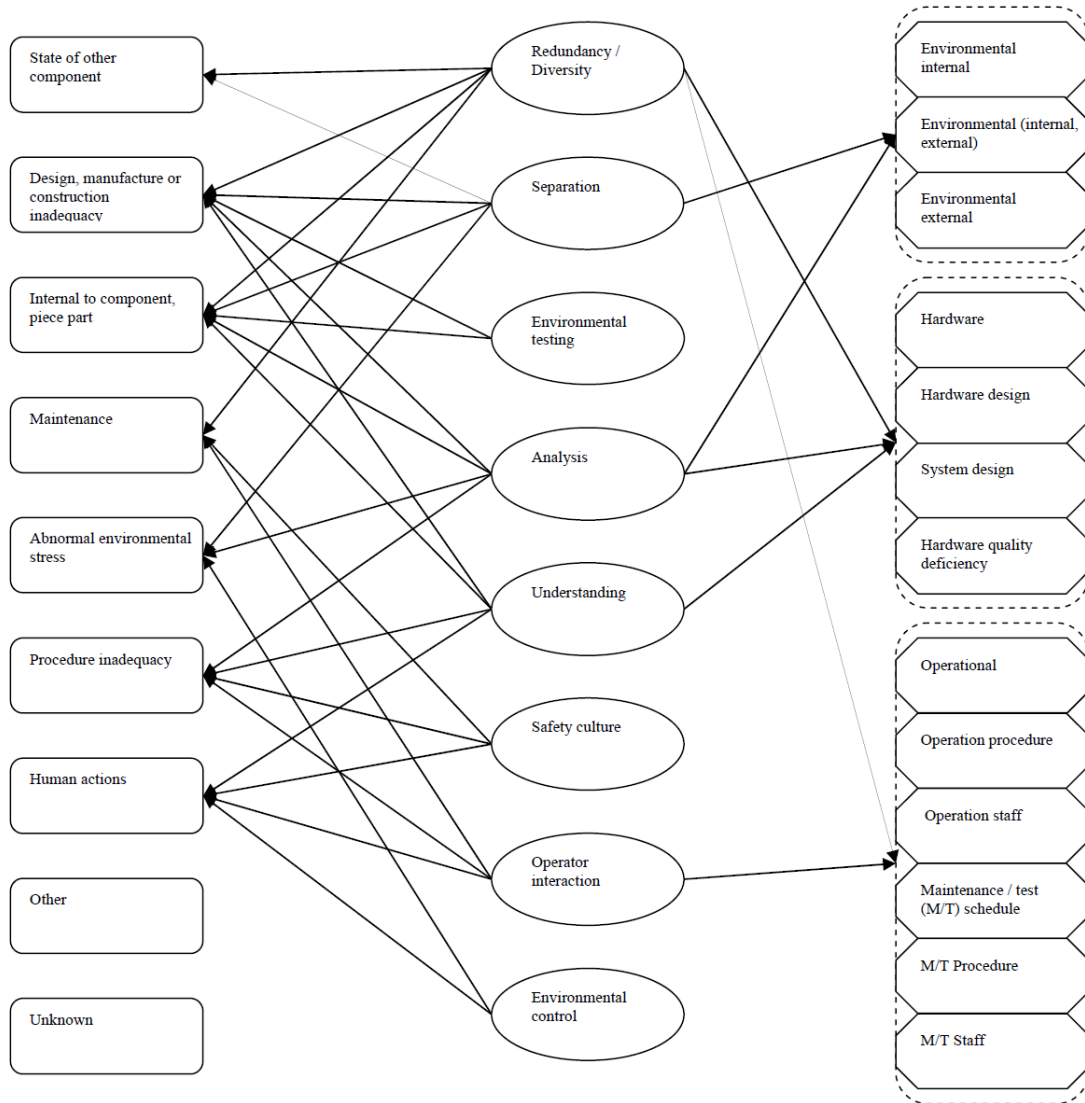


Figure 23: RDR Diagram(Lindberg 2007, p.31)

4.2. Analysis of NRC Common Cause Failure Database

This section will compare the NRC CCFDB against NUREG/CR-6268 Rev 1, which provides the guidance on the classification of failure event data for inclusion in the

NRC CCFDB. The suitability of the failure classification taxonomy will be assessed for suitability to establish a one to one relationship between failure causes and defenses.

#### ***4.2.1. CCFDB Classification System***

In order to interpret the CCFDB classification system, definitions for the terms used may be required. Appendix 2 has a reproduction of the classification definitions from NUREG/CR-6268 Rev 1.

A comparison of the NUREG/CR-6268 Rev 1 and CCFDB failure cause taxonomy is contained in Table 14. A comparison of the NUREG/CR-6268 Rev 1 and CCFDB coupling factor taxonomy is contained in Table 15.

The coupling factor taxonomy in the CCFDB is exactly the same as the NUREG. However the failure cause taxonomy has large differences. This highlights the difficulties with differences, not just between taxonomies, but also relevant to the same database evolving over time.

**Table 14: Comparison of NUREF/CR-6268 Failure Cause Classification and the CCFDB**

<b>Failure Cause Classification</b>		
<b>NUREG/CR-6268</b>	<b>CCFDB</b>	
<b>Design</b>		
Installation / Construction Error	Construction/installation error or inadequacy	DC
Design Error	Design error or inadequacy	DE
Manufacturing Error	Manufacturing error or inadequacy	DM
Design Modified Error		
<b>Operations / Human</b>		
Accidental Action	Accidental human action	HA
Inadequate/Incorrect Procedure	Inadequate maintenance	HM
Failure to Follow Procedure	Human action procedure	HP
Inadequate Training		
Inadequate Maintenance		
<b>External Environment</b>		
Fire/Smoke	Ambient environmental stress	EA
Humidity/Moisture	State of other component	EC
High/Low Temperature	Extreme environmental stress	EE
Electromagnetic Field		
Radiation		
Bio-organisms		
Contamination/Dust/Dirt		
Acts of Nature (Wind/Flood/Lightning/Snow/Ice)		
<b>Internal to Component</b>		
Normal Wear	Internal to component, piece-part	IC
Internal Environment	Internal environment	IE
Early Failure	Setpoint drift	IQ
	Age/Wear	IW
<b>Miscellaneous</b>		
State of Other Component		
Unknown	Unknown	OK
Other	Other	OT
	Inadequate procedure	PA

**Table 15: Comparison of NUREF/CR-6268 Coupling Factor Classification and the CCFDB**

<b>Coupling Factor Classification</b>		
<b>NUREG/CR-6268</b>	<b>CCFDB</b>	
<b>Environmental</b>		
Enviro External	Enviro External	EE
Enviro Internal	Enviro Internal	EI
<b>Hardware Design</b>		
Hardware Design Parts	Hardware Design Parts	HDCP
Hardware Design System	Hardware Design System	HDSC
<b>Hardware Quality</b>		
Hardware Quality Install	Hardware Quality Install	HQIC
Hardware Quality Manufacturing	Hardware Quality Manufacturing	HQMM
<b>Operations Maintenance</b>		
Ops Maint Schedule	Ops Maint Schedule	OMTC
Ops Maint Procedure	Ops Maint Procedure	OMTP
Ops Maint Staff	Ops Maint Staff	OMTS
<b>Operations Operational</b>		
Ops Ops Procedure	Ops Ops Procedure	OOOP
Ops Ops Staff	Ops Ops Staff	OOOS

**4.2.2. Analysis of Observed CCF Events**

In order to provide a qualitative analysis of the CCFDB failure taxonomy, the observed CCF events are analyzed to determine if there is a correlation between the failure cause and the coupling factor. The distribution of CCF events is shown in Table 16.

Note the colors used in Table 16 will be used to discuss a proposal for reclassification of the CCFDB in section 4.3.

**Table 16: Comparison of failure cause and coupling factor for observed CCF events in the CCFDB**

Failure Cause	Coupling Factors										
	Enviro External	Enviro Internal	Hardware Design Parts	Hardware Design System	Hardware Quality Install	Hardware Quality Manufact	Ops Maint Schedule	Ops Maint Procedure	Ops Maint Staff	Ops Ops Proc	Ops Ops Staff
	EE	EI	HDCP	HDSC	HQIC	HQMM	OMTC	OMTP	OMTS	OOOP	OOOS
'Construction/installation error inadequacy'	0	0	63	2	4	0	0	3	4	0	0
'Design error or inadequacy'	3	2	38	7	2	0	1	3	0	1	1
'Manufacturing error or inadequacy'	0	0	6	0	1	3	0	0	2	0	0
'Ambient environmental stress'	3	2	0	0	0	0	1	0	0	1	0
'State of other component'	2	0	4	13	0	0	1	0	0	0	0
'Extreme environmental stress'	32	9	4	2	0	0	2	0	0	0	0
'Accidental human action'	0	0	1	0	1	0	0	4	8	0	0
'Inadequate maintenance'	2	1	1	1	0	0	3	2	0	0	0
'Human action procedure'	0	0	2	0	0	0	0	3	3	1	0
'Internal to component, piece-part'	0	4	17	4	0	3	43	9	0	0	0
'Internal environment'	4	11	0	0	0	0	0	2	0	0	0
'Setpoint drift'	0	0	1	0	0	0	0	1	0	0	0
'Age/Wear'	1	1	3	1	0	1	3	0	0	0	0
'Unknown'	0	0	1	0	0	0	0	0	0	0	0
'Other'	0	0	4	2	0	0	1	0	1	0	0
'Inadequate procedure'	0	0	2	0	0	0	1	24	7	1	0

Table 16 shows that for each cause there is a reasonable spread of coupling factors through which the cause can propagate. Likewise for each coupling factor there is a number of causes. For example, if a failure occurs due to design error, this data shows that there are 9 possible coupling factors that failure could propagate through. For another example, if two components share a maintenance procedure, there are 8 types of failure causes which could affect both components.

This result suggests that there is no suitable relationship between the failure cause and coupling factor. Some classifications do not seem possible, such as construction errors propagating through an operator. The construction error may be related to how the person uses the component, and they made the same error on the other component. What is the coupling factor which propagates the failure? If the operator also used another component of different type, it is unlikely the error would reoccur. If a different operator used the component with the construction error, would the same error have reoccurred, likely. So it is not the human operator which propagates the failure, it's any component that shares the same construction.

With a new requirement to link the coupling factor and failure cause together, it is evident that the classifications for the CCFDB would require redefinition to allow a cause based CCF model to use the data.

#### **4.3. Failure Cause and Coupling Factor Taxonomy Proposal**

The previous section identified the need for a revised failure cause and coupling factor taxonomy. It should be noted that the purpose of this thesis is not to define the categories for a new taxonomy. This must be done in consultation with industry experts. Instead the proposal here is to identify the features of a suitable taxonomy and demonstrate what a proposed category would look like.

In order for such a taxonomy to meet the needs of the proposed cause based models, the following attributes are desired.

- Allow the coupling factor that a failure could propagate through to be inferred from the failure case.
- Minimize ambiguity between definitions.
- Not too complicated.
- Allow the coupling factors between components to be assessed through a qualitative assessment.
- Align to a method of quantitative modeling.

An assessment has been made of the current taxonomy for attributes which meet these criteria. Table 16 contains green cells for failure cause categories and coupling factors which are suitable to keep. Yellow cells will be required to combine or divide into other categories. Red cells are proposed to be removed from the taxonomy.



The coupling factor categories were assessed as being excellent categories which have little ambiguity, and may easily be assessed through a qualitative assessment. Therefore a corresponding failure cause category must be identified. Where a suitable failure cause category could be directly classified under the new taxonomy, a green cell is marked.

Examples of yellow categories to be amended include accidental human action and human action procedure, where it's not clear what coupling factor from a qualitative assessment would propagate these failures. If they were reclassified into either installation human error, maintenance human error, or operations human error it would match the coupling factors.

A proposed taxonomy that achieves the desired features may look like Table 17, There is a direct relationship between the coupling factor and the failure cause. The list is manageable, with training the definitions could be unambiguous.

**Table 17: Proposed Cause and Coupling Factor Taxonomy**

Failure Cause	Coupling Factor
Installation Procedural Cause	Same Install Procedure
Installation Human Cause	Same Install Team
Component Design Deficiency	Same Component Design
System Design Deficiency	Same System Design
Age/Wear	Same Age within Mission Period
Component Manufacturer Fault	Same Component and Manufacturer
Operator Error	Same Operators

Operation Procedure Error	Same Operating Procedures
Maintainer Error	Same Maintenance Team
Maintenance Procedure Error	Same Maintenance Procedure
Maintenance Schedule Error	Same Maintenance Schedule
Environment Internal Induced	Same Fluid
Environment External Induced	Same Location

#### **4.4. Summary**

An enabler for the success of the proposed CCF models is a failure event taxonomy which allows for inference to be made about the possible propagation means for a failure knowing only the failure cause.

An assessment was conducted of the CCFDB and it was found that there is insufficient correlation between the failure cause and coupling factor categories for it to be used in its current form.

The features of a suitable taxonomy system were described and a failure event taxonomy was proposed based on the existing coupling factor categories, for review by industry experts.

## Chapter 5: Existing CCF models

### ***5.1. Introduction***

The phenomena of Common Cause Failure has been recognized as a consideration in design for some time. It was first discussed as a specific activity requiring special treatment in the early 1970s (Smith & Watson 1980). The first major use of a CCF failure model was in WASH 1400 probabilistic risk assessment of Nuclear Power Plant safety in 1975 (Rasmussen 1975). Since WASH 1400, over 30 different Common Cause Failure models have been proposed. A complete literature review of these models is included as appendix 1.

This chapter will summarize the models which have been used to propose the model extensions discussed in Chapter 6 and 7. Each model will be described and reviewed for:

- the principles for which the model has been created,
- an opinion on the advantages and limitations of each model, and
- a list of references which have informed the analysis.

## 5.2. Direct Estimates

Direct estimate models consist of:

- Direct Assessment (Qualitative)
- Basic Parameter Model

### 5.2.1. *Direct Assessment (Qualitative)*

The direct assessment model can be considered as a procedure rather than a modeling technique. (Hirschberg 1985) discussed this approach in detail. It involves using the actual number of demands and the number of observed failures with multiplicity,  $i$ , and estimating the quantities of interest directly from the data set. For the purposes of this model  $i$  is defined as any positive integer, and the quantities of interest for this case is defined as the common cause failure rate or common cause failure probability.

This approach is typically simple and is less dependent upon sound knowledge of any mathematical or statistical skills. (Anude 1994)

Advantages include:

- The method is simple as there is minimal data required and minimal mathematical knowledge required to determine broad estimates.
- The model is predicated on experience.

Limitations include:

- The model cannot estimate common cause failures for K out of N events for which it does not have data.

- Component symmetry is assumed (ie.  $X_{AB} = X_{BC}$ )
- Does not allow for partial failure or component degradation
- No inference can be made given knowledge of the failure cause.
- The model cannot account for unique system architecture which may contribute or defend against CCF.

### 5.2.2. Basic Parameter Model

The basic parameter model was proposed by Fleming, et al. in 1983 (Fleming et al. 1983) and calculates the CCF basic event directly from the data. This estimation is given by: (Mosleh et al. 1998)(Mosleh 1991)

$$Q_k^m = \frac{n_k}{N_k}$$

$Q_k^{(m)}$  = *basic event failure frequency/probability for k components failing within a common cause component group of size m, ( $1 \leq k \leq m$ ).*

$n_k$  = *the number of failure events which resulted in k components failing within a common cause component group of size m, ( $1 \leq k \leq m$ ).*

$N_k$  = *the number of demands on any k component in the common cause group.*

If it is assumed that each time the system is operated, all of the m components in the group are demanded, then. <sup>8</sup>:

---

<sup>8</sup> This estimator can change depending on the scheme used to test components. This estimator is for non-staggered testing. Other estimators and discussion on testing schemes is provided in NUREG /CR-5485 (Mosleh et al. 1998)

$$N_k = \binom{m}{k} N_D$$

$$Q_k^m = \frac{n_k}{\binom{m}{k} N_D} \quad \text{non-staggered testing}$$

$$Q_k^m = \frac{n_k}{m \binom{m}{k} N_D} \quad \text{staggered testing}$$

$N_D =$  the number of demands on the system (or time  $T$ )

The total component failure rate can be calculated as the sum of the CCBEs:

$$Q_t = \sum_{k=1}^m \binom{m-1}{k-1} Q_k^m$$

Replacing  $Q_k^{(m)}$  with its estimator gives the following estimator for the total failure rate,  $Q_t$ :

$$Q_t = \frac{1}{m N_D} \sum_{k=1}^m k n_k \quad \text{non-staggered testing}$$

$$Q_t = \frac{1}{m^2 N_D} \sum_{k=1}^m k n_k \quad \text{staggered testing}$$

- Advantages include:
- The method is simple as there are no intermediate steps in quantifying basic common cause events.
- The model is intuitive.
- Suitable for any amount of redundancy (for which data is available).
- No need to differentiate between independent and common cause failures.

Limitations include:

- The model cannot estimate common cause failures for redundancy configurations for which data is unavailable.
- Component symmetry is assumed (ie.  $X_{AB} = X_{BC}$ )
- Does not allow for partial failure or component degradation.

### 5.3. Ratio Models

Ratio models are based on the hypothesis that system specific estimates for CCF can be made by combining generic average ratio parameters with system specific single/total failure rates (Vaurio 2008). This provides the advantage that ratio models can be estimated from specific data collection activities such as the Common Cause Failure Data Base and applied to areas where CCF data may not exist.

Ratio models have the following advantages:

- There is a direct and intuitive quantity to the model parameters.
- Generic ratio parameters can be calculated from generic data and then applied to plant specific single failure rates. This reduces the data requirements compared to direct estimate models.
- Success data is not required to estimate the model parameters.

The limitations of all ratio models discussed here include:

- The model assumes a transferable empirical ratio between failure rates and Common Cause Failure rate.
- The ratio models described here assume component symmetry.

- No inference can be made given knowledge of the failure cause.
- The model cannot account for unique system architectures which may contribute or defend against CCF.
- Confusion in the interpretation of single failures being modeled as independent failures, particularly when applying impact mapping rules.

### 5.3.1. Beta Factor Model

The Beta Factor Model, proposed by Fleming in 1975 (Fleming 1975), is a component failure ratio model which is one of the most popular where generic data used to estimate parameters are limited. It is still the most commonly used CCF model outside the nuclear industry (Hokstad & Rausand 2008).

The basic parameters can be calculated as (Mosleh et al. 1998):

$$Q_k^{(m)} = \begin{cases} (1 - \beta)Q_t & k = 1 \\ 0 & m > k > 1 \\ \beta Q_t & k = m \end{cases}$$

$Q_t$  = The total failure probability of one component

$m$  = Common Cause Component Group size

$k$  = Number of failed components due to common cause failure

The MLE parameter estimate is (Mosleh et al. 1998)<sup>9</sup>:

---

<sup>9</sup> This estimator can change depending on the scheme used to test components. This estimator is for non-



$$\hat{\beta} = \frac{\sum_{k=2}^m k n_k}{\sum_{k=1}^m k n_k}$$

$n_k$  = number of events involving  $k$  components in a failed state

$m$  = the number of components within the CCCG

$\beta$  (usually between 0.01 and 0.3) (Anude 1994) is defined as the point estimate of the conditional probability that a unit failure is a Common Cause type. The Beta Factor model uses one parameter in addition to the total component failure probability to calculate the Common Cause failure probabilities regardless of the size of the Common Cause Component Group.

The advantages of the Beta Factor model, in addition to the ratio model advantages are:

- Simplicity compared to other ratio models.
- Regardless of the number of components comprising the system, it requires the estimation of only two parameters.

The limitations of the Beta Factor model, in addition to the ratio model limitations are:

- The model does not acknowledge CCFs of various multiplicities within the Common Cause Component group. Failure can either be one component or the whole component group. (Hokstad 2004)

---

staggered testing. Other estimators and discussion on testing schemes is provided in NUREG /CR-5485 (Mosleh et al. 1998)

- For most redundant systems, this model has been proven to be excessively conservative and pessimistic in predicting CCF failure rates. (Mosleh et al. 1998)
- Rigorous estimators for the beta factor model parameters are fairly difficult to obtain, although approximate methods have been developed and used in practice.(Mosleh et al. 1998) (Mosleh 1986)

### ***1.1.1. Partial Beta Factor (PBF) Model***

The Partial Beta Factor (PBF) model was first conceived by Edwards in 1982 and later developed by Johnston (Johnston 1987) to allow consideration for the target system dependencies and defenses. After a qualitative analysis identifies CCCGs containing identical components and a criticality assessment of the effect from dependencies based on cut sets, a matrix is created allowing the different attributes leading to dependencies between the components to be evaluated. A Beta Factor is then created as a product of a number of partial beta derived from judgments of system defenses.

$$\hat{\beta} = \prod_j \beta_j$$

$\beta$  = *The beta factor for the Beta Factor model.*

$\beta_j$  = *The partial beta factors for defence j attribute*

The basic parameters are the same as the beta factor model and calculated as:

$$Q_k^{(m)} = \begin{cases} (1 - \beta)Q_t & k = 1 \\ 0 & m > k > 1 \\ \beta Q_t & k = m \end{cases}$$

A limitation of this approach is that each partial beta factor multiplies the whole failure rate for defenses which may only affect a portion of the failure rate. Johnson proposed an extension to the PBF model where the failure rate is separated into specific causes and the Partial Beta Factor only adjusts that portion of the failure rate (Johnston 1987).

$$Q_{m,i}^{(m)} = Q_i \prod_j \beta_{j,i}, \quad Q_m^{(m)} = \sum_i Q_{m,i}^{(m)}$$

$Q_i$  = The failure probability/rate of cause  $i$ .

$\beta_{j,i}$  = The partial beta factor for defence  $j$  and cause  $i$

It should be noted that the outcome of the Partial Beta Factor model is to arrive at a system specific Beta Factor, and as such multiplicity of failures within a common cause component group is not recognized.

### **1.1.2. Alpha Factor Model (AFM)**

The alpha factor model (AFM) is a failure event ratio model that was first proposed by Mosleh and Siu in 1987 (Mosleh & Siu 1987). Each  $\alpha_k$  factor is the probability that given a failure it will fail  $k$  components out of  $m$  components within the CCCG. The AFM parameters are defined and calculated as (Mosleh et al. 1998):

$$\alpha_k = \frac{n_k}{\sum_{k=1}^m n_k}$$

$m$  = the number of redundant components

$n_k$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$\alpha_k$  = the fraction of total failure events/frequency that occur in the system resulting in  $k$  out of  $m$  failures.

The alpha factor method is used to estimate the basic event probabilities using (Mosleh et al. 1998) .

$$Q_k^{(m)} = k \binom{m-1}{k-1}^{-1} \cdot \frac{\alpha_k}{\alpha_t} \cdot Q_t \quad \text{non-staggered test data}$$

$$Q_k^{(m)} = \binom{m-1}{k-1}^{-1} \cdot \alpha_k \cdot Q_t \quad \text{staggered test data}$$

where:

$$\binom{m-1}{k-1} = \frac{(m-1)!}{(k-1)! (m-k)!} \quad \text{and} \quad \alpha_t = \sum_{i=1}^m k \alpha_k$$

$Q_k^{(m)}$  = *basic event failure frequency/probability for k components failing within a common cause component group of size m, (1 ≤ k ≤ m).*

$Q_t$  = *total failure frequency/probability of each component due to independent and common cause events.*

This formulation has the property that that  $\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_m = 1$  so that the  $\alpha$ s are not mutually independent. (Vaurio 2008).

Due to AFM's ability to calculate its parameters directly from the impact vectors, the AFM is sometimes used as an intermediate step in calculating Beta Factor and MGL parameters. The relationship between these parameters is detailed for different systems within NUREG/CR-5485 (Mosleh et al. 1998).

Advantages over the Beta Factor model are that the AFM model can model various

multiplicities of failure within the Common Cause Component Group and unlike the Beta and MGL methods, AFM's parameters are directly related to measurable properties of the system and are usually calculated directly from observed data as impact vectors (Mosleh 1991).

#### 5.4. Shock Models

Shock models are based on the hypothesis that each component within the CCCG undergoes shocks according to a Poisson process. For each component within the CCCG the shock is a Bernoulli trial which will fail the component with probability  $\rho$ .

Most shock models are adoptions or simplifications of the multivariate exponential model derived by Marshall and Olkin in 1967 (Marshall & Olkin 1967). For these models the number of failed components,  $k$ , resulting from a shock is binomially distributed. Shock models strive to model the actual physical phenomena that result in CCF to occur.

Shock models have the following advantages:

- Can be used to model high levels of redundancy.(Anude 1994)
- Can estimate CCF frequency even when CCF events have not been observed.  
(Atwood 1986)
- Easier to adjust for different sizes CCCG groups. (Kvam & Martz 1995)

- Importing/exporting data for different sized systems is more accurate and often easier due to the ability to characterize the underlying probability of common cause failures. (Kvam & Martz 1995)

Shock models have the following disadvantages:

- Includes parameters which are difficult to measure with data (such as shock rates).
- Requires demand/success data to calculate parameters.
- Confusion in the interpretation of single failures being modeled as independent failures, particularly when applying impact mapping rules.
- Lethal shocks need to be distinguished from multiple CCF failing all system components.
- Assumes component symmetry (ie.  $X_{AB} = X_{BC}$ ).
- Assumes that given a shock has occurred, items will fail independently which may be violated in practice. (Anude 1994)
- Assumes zero time to repair. (Atwood 1986)
- Assumes renewal to as good as new. (Atwood 1986)
- Assumes Constant Failure Rates. (Atwood 1986)
- The  $\rho$  parameter is independent of the size of the CCCG. (Vaurio 1999)
- Any subset of  $k$  components of a system of size  $m$  is equally vulnerable to exactly the same common-causes and stresses as in a system of size  $k$ , or

anything larger than  $k$ . This results in the assumption that  $n_k > n_{k+1}$  (the mapping rule). (Vaurio 1999)

- Data is needed from a system with at least  $m=3$  in order to solve the three unknowns. (Vaurio 2008)
- The analyst needs to distinguish between a single CCF and a single independent failure. This can become subjective from fault reports leading to higher uncertainty.
- No parameters are directly linked to the degree of system protection against CCFs.
- Probability that the value of the binomial parameter  $\rho$  remains fixed across all system shocks despite each shock having different intensities and different sources. (Anude 1994)
- Does not model different intensity shocks to the system. (Anude 1994)
- Parameter calculation can be cumbersome. (Kvam 1993)
- No inference can be made given knowledge of the failure cause.
- The model cannot account for unique system architecture which may contribute or defend against CCF.

#### ***5.4.1. Binomial Failure Rate Model***

The Binomial Failure Rate Model (BFRM) model was proposed by Vesely in 1977 (Vesely 1977) to adapt the shock model proposed by Marshall and Olkin. This model was motivated by estimation with less data than previously required and to describe the

underlying failure process generated by CCF events. It assumes that CCF occur when all  $m$  redundant components of a system are challenged by a shock at a rate of  $\mu$ . The number of resulting failures from each shock,  $k$ , is random with a binomial distribution with probability  $\rho$ .

This model has also been known as the three-parameter BFR model with parameters,  $Q_I$  (or  $\lambda$ ),  $\mu$  and  $\rho$ . These parameters can be estimated using (Marshall et al. 1998):

$$Q_I = \frac{n_I}{mN_D}$$

$$\sum_{k=1}^m kn_k = \rho \frac{m \cdot n_t}{1 - (1 - \rho)^m} \quad \text{solve for } \rho$$

$$\mu = \frac{n_t}{N_D} \cdot \frac{1}{1 - (1 - \rho)^m}$$

where

$$n_t = \sum_{i=1}^m n_k$$

$n_k$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$n_I$  = the number of failure events/frequency which resulted in the independent failure of the component.

$n_t$  = total number of common cause failures.

$N_D$  = the number of demands on the system (or time  $T$ ), can also be called  $N_S$

The basic parameters can be calculated as (Vesely 1977):



$$Q_k^m = \begin{cases} Q_I + \mu \cdot \rho(1 - \rho)^{m-1} & \text{where } k = 1 \\ \mu \cdot \rho^k(1 - \rho)^{m-k} & \text{where } 2 \leq k \leq m \end{cases}$$

$Q_I$  = the independent failure rate of each component

$Q_k^m$  = basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$\mu$  = rate of shocks

$\rho$  = probability of component failure given a shock

The rate of failure of  $k$  components is simply the binomial probability of  $k$  in  $m$  components failing multiplied by the rate of shocks. The rate of failure for a single component is the independent failure rate plus the contribution of 1 component failing due to a common cause shock. Probability that the value of the binomial parameter  $\rho$  remains fixed across all system shocks.

Due to its inaccuracy to real systems, the model presented here is rarely used (Mosleh et al. 1988) (Kvam 1998b); instead, a simple binomial shock model using the BFR model with lethal shocks is typically preferred. Note that the BFRM and  $\beta$ -factor model are the same for a two component system. (Rausand & Høyland 2003)

#### **5.4.2. Binomial with Lethal Shocks**

Atwood proposed an extension to the BFR model in 1986 that included an additional independent process of lethal shocks (Atwood 1986). In this model, each lethal shock will fail all components of the system at a rate of  $\omega$ .

This model has also been known as the four-parameter BFR model with parameters,  $Q_I$  (or  $\lambda$ ),  $\mu$ ,  $\rho$  and  $\omega$ . These parameters can be calculated using [85]:

$$Q_I = \frac{n_I}{mN_D}$$

$$\sum_{k=1}^m kn_k = \rho \frac{m \cdot n_t}{1 - (1 - \rho)^m} \quad \text{solve for } \rho$$

$$\mu = \frac{n_t}{N_D} \cdot \frac{1}{1 - (1 - \rho)^m}$$

$$\omega = \frac{n_L}{N_D}$$

where

$$n_t = \sum_{i=1}^m n_k$$

$n_k$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$n_I$  = the number of failure events/frequency which resulted in the independent failure of the component.

$n_L$  = total number of lethal common cause failures.

$n_t$  = total number of common cause failures.

$N_D$  = the number of demands on the system (or time  $T$ ), can also be called  $N_S$

The basic parameters can be calculated as (Atwood 1986):

$$Q_k^m = \begin{cases} Q_I + \mu \cdot \rho(1 - \rho)^{m-1} & \text{where } k = 1 \\ \mu \cdot \rho^k(1 - \rho)^{m-k} & \text{where } 2 \leq k < m \\ \mu \cdot \rho^m + \omega & \text{where } k = m \end{cases}$$

$Q_I$  = the independent failure rate of each component

$Q_k^m$  = basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

- $\mu$  =rate of shocks
- $\rho$  =probability of component failure given a shock
- $\omega$  = rate of lethal shocks

This extension has been found to be more accurate than the basic BFR model (Mosleh et al. 1988). The probability that the value of the binomial parameter  $\rho$  remains fixed across all system shocks.

### 5.5. Interference Models

Interference models also attempt to model the physical phenomena of CCF but without the shock model's assumption of independence. Instead these models predict the number of failures by assuming random variables for component strength and load. When the load exceeds the strength a component is expected to fail. The more intense the load or the more depleted the strength then the higher the probability of failure. There is no explicit distinction between an independent and common cause failure event.

Inference models have the following advantages:

- Can be used to model high levels of redundancy.
- Can estimate CCF frequency even when CCF events have not been observed.
- Easier to adjust for different sizes CCCG groups.
- Models different intensities of shocks to the system.

- Directly models the system's protection against CCF through the resistance measure.
- There is no need to distinguish between a single CCF and a single independent failure.
- Lethal shocks are quantified by their shock intensity and included within the model formulation.
- Importing/exporting data for different sized systems is more accurate and often easier due to the ability to characterize the underlying probability of common cause failures. (Kvam & Martz 1995)

Interference models have the following disadvantages:

- Requires a probability distribution to be estimated for shock and resistance intensities. This requires knowledge of the physical characteristics of the components and the data required to quantify distributions differs from just failure and success data.
- Requires demand/success data to calculate parameters.
- Assumes component symmetry (ie.  $X_{AB} = X_{BC}$ ).
- Assumes zero time to repair.
- Assumes renewal to as good as new.
- Assumes Constant Failure Rates.

- Any subset of  $k$  components of a system of size  $m$  is equally vulnerable to exactly the same common-causes and stresses as in a system of size  $k$ , or anything larger than  $k$ . This results in the assumption that  $n_k > n_{k+1}$  (the mapping rule). (Vaurio 1999)
- No inference can be made given knowledge of the failure cause.
- The model does not explicitly account for unique system architecture which may contribute or defend against dependencies between components.

### *5.5.1. Common Load Model*

The Common Load model proposed by Mankamo and Kosonen in 1977 (Mankamo 1977) is based on a load-strength interference methodology for describing the failure mechanism. The model interprets the failure mechanism as a load imposed on a component where the components strength is tested. A failure occurs when the resistance is not sufficient to withstand the load.

When it comes to redundant systems of components, the load posed to the system is shared by all the components of the system equally, and a failure of certain multiplicity is determined by the number of components whose resistance is exceeded by the load. Both the load and the component resistance are described in terms of random variables and assumed probability distributions (Zitrou 2006b).

The probability density function of the resistance,  $R$ , is denoted by  $f_R(x)$ . In the event

of an occurrence of a random shock,  $S$ , with a probability density function of  $g_s(x)$ , then the event of having exactly  $k$  of the components fail simultaneously, is given as (Anude 1994):

$$Q_k^{(m)} = P(R_k \leq S < R_{k+1})$$

$$Q_k^{(m)} = \int_0^{\infty} \frac{m!}{k!(m-k)!} (F_R(y))^k (1 - F_R(y))^{m-k} g_s(y) dy$$

$S$  = the random variable for the shock intensity

$R_k$  = the random variable for resistance intensity where

$$R_1 \leq R_2 \leq \dots \leq R_n$$

$g_s(x)$  = the probability distribution for the shock random variable

$F_R(x)$  = the cumulative probability distribution for the resistance random variable

$k$  = the multiplicity of failure being investigated

$m$  = the number of components within the CCCG

The model has a fixed number of parameters, independent of the size of the system. Like the Shock Models, the model can be applied to any failure multiplicities. The model assumes that the  $n$  components of a system have independent and identically distributed random resistances  $R_1, R_2 \dots R_n$ . (Anude 1994)

Cases of non-symmetry can be modeled by removing the assumption of identical distributed components and creating separate  $f_R(x)$  distributions for each component.

## 5.6. Other Models

### 5.6.1. Reliability Cut Off Method

The Reliability Cut Off Method was proposed by Bourne et. al in 1981 (Bourne et al.

1981) as a system level estimate of CCFs based on an assessment of the vulnerability of the system. No identification of CCCG is conducted and the methodology assumes that the unreliability of a system due to CCFs can never exceed some limiting values, determined by system design. These estimates do not involve the use of data and instead use generic estimates from experts.

The original article for this model could not be obtained and so a full assessment could not be completed.

### ***5.6.2. Unified Partial Method***

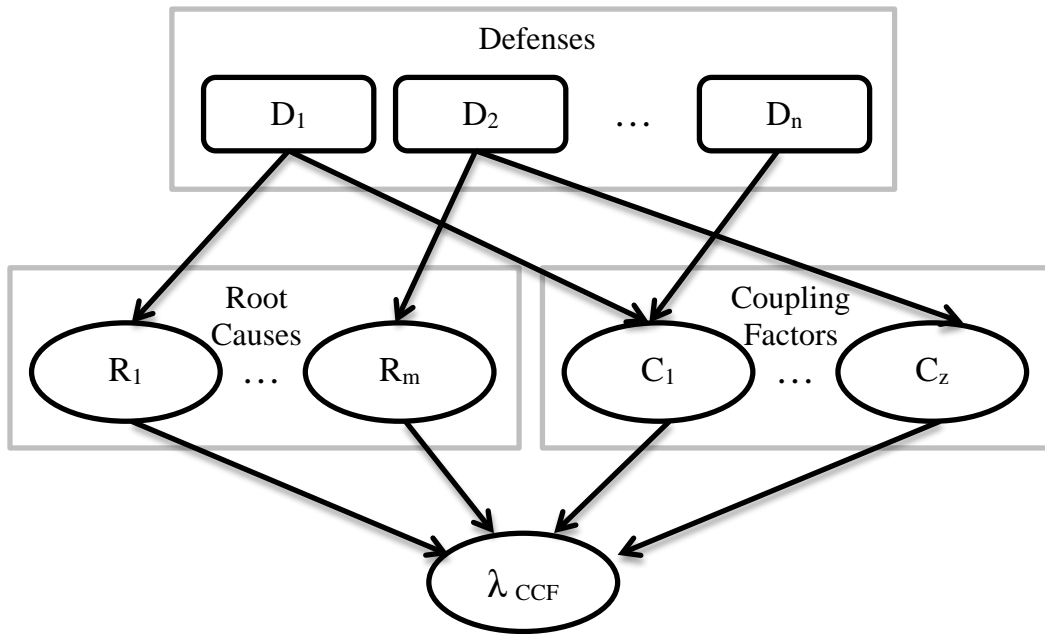
The Unified Partial Method (UPM) (Brand & Gabbot 1993) is the current method which has been adopted by the UK nuclear industry. UPM is a methodology to assess the vulnerability of a system to CCF and uses one of two models to quantify its estimates, the Partial Beta Factor method for component level analysis, and the Cut-Off method for system level analysis. Brand describes UPM as not being a complete method for dependent failure assessment, but a useful methodology for ‘standard systems’ (Mosleh et al. 1998).

### ***5.6.3. Influence Diagram Model (Zitrou 2006a)***

Zitrou in 2006 proposed an extension of the UPM model using influence diagrams and a more detailed mathematic formulation using Bayesian methods. The objective of

Zitrou's research was to explore the modeling of CCF using advanced mathematical techniques (influence diagrams). Zitrou wanted to keep the desirable features of UPM where attributes of the system are included in the model, the ability to provide estimates in the absence of data and the simplistic application of the method by analysts. Zitrou wanted to use the influence diagram to extend UPM's accuracy by modeling the dependency between defenses and improve the models quantitative estimates.

Zitrou's model consisted of the creation of an influence diagram which followed the convention of figure 5. (Zitrou 2006a, p.18)



**Figure 24: Zitrou General Influence Diagram Structure**

The specific taxonomy used to define the ID nodes were the same as for UPM. The specific dependencies between nodes were established using an expert elicitation



technique.

Two unique elements are proposed in Zitrou's model (Zitrou 2006a, p.257).

- The definitions of the dependencies between defenses were established to determine if improving one defense would have a positive, negative or neutral effect on another defense.
- A geometric scaling model was proposed which is used to quantify the effect of the defense levels on the probability of root causes and coupling factors. This model reduces the burden of the quantification process by allowing the root cause and coupling factor probability distributions to be determined based on a base defense level. The geometric scaling model can then scale the probability distributions dependent on the level of defense applicable.

Zitrou's model achieves the following objectives (Zitrou 2006a, p.254):

- Incorporates the qualitative advantages of the UPM model.
- After quantification by experts the model can be easily used by practitioners.
- Extends the casual modeling of UPM to a finer level.
- Captures the dependency between defenses.
- Captures the uncertainty of the expert judgment.
- Provides an investigative framework in which conditional probabilities can be explored.

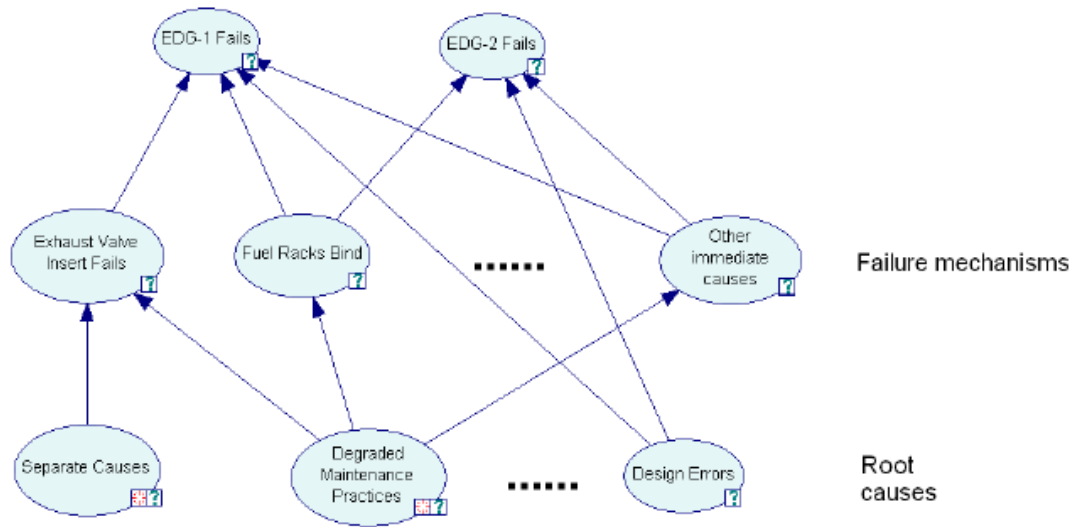
Zitrou's research goal was to conduct an exploration of using influence diagrams to model CCF. The thesis proposed a methodology and conducted the limited development of a quantification model on Emergency Diesel Generators using some expert estimation. The example model was not fully developed and verification against known system results was not conducted. The model did not consider incorporating data analysis techniques from CCF databases.

#### ***5.6.4. CCF model for Event Assessment (Kelly et al. 2011)***

Kelly et al., have written a draft paper which demonstrates the inability of current models to conduct event assessments. The paper proposes using a Bayesian Network to model the causal relationship between root causes, failure mechanisms and the CCF event probability. Two methods of constructing this model are proposed:

- The first model explicitly models the root causes and failure mechanisms specific to the component.
- The second model uses the generic CCF taxonomy used by the INL CCF database.

The paper focuses on conducting event assessments where a failure cause is known; as such the model does not include the coupling factors or defenses. An example of such a model is included as Figure 66. This paper expresses an ideology for CCF modeling but does not propose specific model construction or quantification details. This paper forms the objective of this research.



**Figure 25: Bayesian network representing more general situation of multiple failure mechanisms and causes in a CCCG of two EDGs (Kelly et al. 2011, p.6)**

### 5.7. Model Comparison

Each model has been assessed against the criteria set in the research objectives which is summarized in Table 18. A description of how these models have contributed to each proposed model will be provided in Chapter 6 and 7.

**Table 18: Assessment of previous CCF Models**

	Basic Parameter	Beta Factor	Partial Beta Factor	Alpha Factor Model	Binomial Failure Rate Model with Lethal Shocks	Common Load	Reliability Cut Off	Influence Diagram	Bayesian Network
<b>Feature Description</b>	<b>BP</b>	<b>BF</b>	<b>PBF</b>	<b>AFM</b>	<b>BFRL</b>	<b>CL</b>	<b>RCO</b>	<b>ID</b>	<b>BN</b>
Explicitly Models System Features	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Models failure cause	N	N	P	N	N	N	P	Y	Y
Models failure cause defense	N	N	Y	N	N	N	Y	N	P
Models coupling factor	N	N	P	N	N	N	P	Y	N
Models coupling factor defense	N	N	Y	N	N	N	Y	N	N
Models deeper causal levels	N	N	N	N	N	N	N	N	Y
Models cause condition / shock	N	N	N	N	Y	Y	N	N	Y
Models multiplicity of failures within CCCG	Y	N	N	Y	Y	Y	N	N	Y
Models includes consideration for rectification period	N	N	N	N	N	N	N	N	N
<b>Common Cause Component Grouping Characteristics</b>	<b>BP</b>	<b>BF</b>	<b>PBF</b>	<b>AFM</b>	<b>BFRL</b>	<b>CL</b>	<b>RCO</b>	<b>ID</b>	<b>BN</b>
Model non-symmetrical but similar components within the same CCCG	N	N	N	N	N	N	N	N	Y
Model different components within the same CCCG	N	N	N	N	N	N	N	N	Y
A component can be part of many CCCGs	N	N	N	N	N	N	N	N	Y
No limit to CCCG size	Y	Y	Y	Y	Y	Y	Y	Y	Y
Model different failure multiplicities within the CCCG ( $k$ failures in $n$ )	Y	N	N	Y	Y	Y	N	N	Y

Event Assessment Capabilities	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Event Assessment with knowledge of a failed component	Y	N	N	Y	Y	?	N	Y	Y
Event Assessment with knowledge of failure cause	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence - Partial Failures	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence- Virtual evidence of cause	N	N	N	N	N	N	N	Y	Y
Parameter Estimation	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Impact Vector Method (including method for incorporating uncertainty)	Y	P	N	Y	Y	N	N	N	N
Expert estimations (in absence of any data)	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account for reliability growth (discount previous failures)	N	N	N	N	N	N	N	N	N
Update parameters with new evidence	Y	P	N	Y	Y	Y	N	N	N
Incorporate evidence from different sized CCCGs	N	P	N	P	Y	Y	N	N	N
Account for CCF which occurred in a different mission time	N	N	N	N	N	N	N	N	N
Account for CCF data which has artificial separation in time.	N	N	N	N	N	N	N	N	N
Use system specific failure rate data combined with generic model parameter	N	Y	N	Y	N	N	N	N	N
Uncertainty Characteristics for Parameter Estimation	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Does not require to distinguish between independent and single CCF failures	Y	Y	Y	Y	N	Y	Y	Y	Y
Failures outside the mission period	Y	P	N	Y	Y	N	N	N	N
Uncertainty of shared cause	Y	P	N	Y	Y	N	N	N	N
Uncertainty of coupling factor	Y	P	N	Y	Y	N	N	N	N
Uncertainty in intervals due to staggered testing	P	P	N	P	P	N	N	N	N
Partial failures and component degradation	Y	P	N	Y	Y	N	N	N	N
Usability and Cultural Considerations	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Backward compatible to Alpha Factor Model parameters	Y	N	N	Y	N	N	N	N	N
The time investment is no more than the alpha factor model.	Y	Y	Y	Y	Y	N	Y	N	N
Automatic parameter estimation is possible from the CCFDB/RADs	Y	Y	N	Y	Y	N	N	N	N

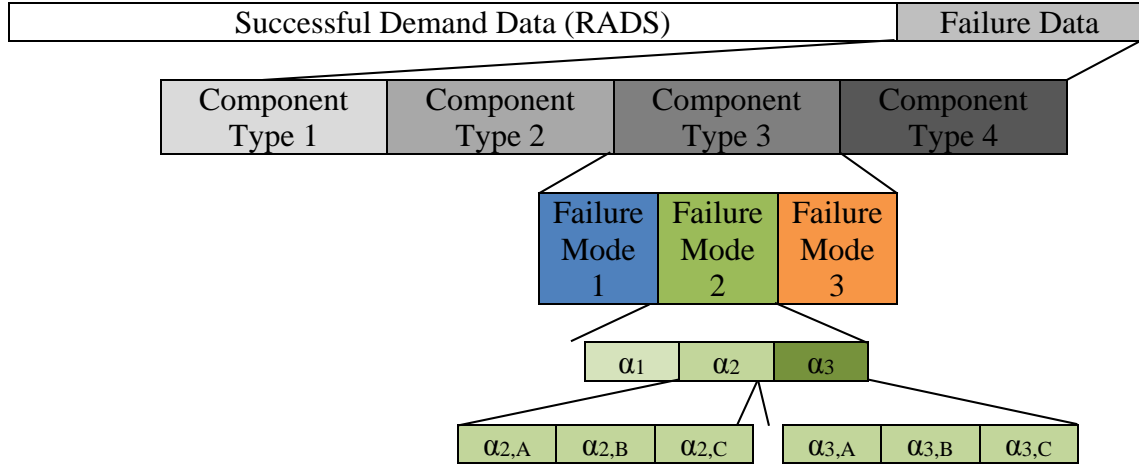
## Chapter 6: Partial Alpha Factor Model (PAFM)

### 6.1. Introduction

#### 6.1.1. *Motivation for PAFM*

Impact vectors and data collection activities such as the ICDE database made quantitative CCF modeling accessible to analysts. However, as shown in Chapter 5, all models which rely on impact vectors, do not use the information about failure causes which exists within these databases. This information has only been used to obtain insights into the CCF phenomena through qualitative analysis (Lindberg 2007; Wierman et al. 2003a; Wierman et al. 2003b; Wierman et al. 2003d; Wierman et al. 2003c). Since 2004 the failure causes for single failure data has also been recorded within the NRC CCFDB (Wierman 2013). This allows enough information to further conditionalize the Alpha Factor Model (AFM) parameters based on failure causes.

The creation of the Partial Alpha Factor Model (PAFM) is motivated by the ability to allow event assessments to be conducted using knowledge of the failure cause, whilst minimizing changes to the already popular AFM methodology which is well understood by the nuclear industry. The AFM parameter estimates are typically conditionalized by component type, failure mode and multiplicity of failure. The PAFM will further conditionalize the alpha factor parameter by failure cause as shown in Figure 26.



**Figure 26: Failure Event Conditionalisation**

This conditionalisation has some limitations when limited data is available, however the system level estimates will be no worse than simple use of the AFM. This conditionalisation has advantages other than event assessment such as the course modeling of asymmetrical dependencies between components. A complete comparison of the PAFM against previously proposed models will be discussed at the end of this chapter.

### **6.1.2. Chapter Scope**

This chapter will discuss:

- An overview of the model
- Description of the model parameters.
- Formulate the parameter estimation equations
- Describe methods to quantifying parameters.

- Describe a system analysis method.
- Describe an event assessment method.
- Describe the data requirements for the model.
- Summarize and evaluate the model.

### **6.1.3. Examples**

In order to demonstrate CCF analysis using the PAFM, the same two examples used in Chapter 2 will be used. Example 1 is a two train EDG example where the system features of the components make them near identical, and example 2 is a mixed redundancy system of two EDGs and three pumps.

### **6.2. Model Overview**

A brief overview of the PAFM's use within the CCF analysis process will be provided to give context to the model parameter development. A more detailed account of how the PAFM is used within the PRA is provided in section 6.6.

The PAFM uses the same CCF methodology as the AFM described in Chapter 2 with two key differences. A component may be a member of multiple CCCGs based on shared coupling factors. The alpha factors for each CCCGs basic events are calculated as a combination of partial alpha factors calculated using impact vectors for each failure cause. Specifically the CCF analysis procedure has the following modifications:



1. *Qualitative Screening* (ref section 2.4.1 on page 24). A component may be a member of multiple CCCGs based on its coupling factor with other components. However, for each coupling factor, the component can only be member of one CCCG. Pump 1 shares its maintenance team and location with pump 2. Pump 1 also shares its installation procedure with pump 3. Pump 1 will be part of two common cause component groups,  $CCCG_X = \{P_1, P_2\}$  and  $CCCG_Y = \{P_1, P_3\}$ .
2. *Identification of Common Cause Basic Events* (ref section 2.6.1 on page 28). The CCCBEs will be constructed with consideration for all CCCGs which the component is a part of. E.g.  $P_1 = P_{1,i} + X_{P_1P_2} + Y_{P_1P_3}$ .
3. *Parameter Representation of CCBEs* (ref section 2.6.3 page 33). The CCBEs are quantified using the Basic Parameter which accounts for multiple CCCGs. For example  $P(P_{1,i}) = Q_1^{(2)}$ ,  $P(X_{P_1P_2}) = Q_2^{(2)[X]}$  and  $P(Y_{AC}) = Q_2^{(2)[Y]}$
4. *Alpha Factor Model Parameterization* (ref section 2.6.4 page 36). The new Basic Parameter values are quantified using a combination of partial alpha factors which will be described within this chapter. E.g.
$$Q_2^{(2)[X]} = \left( \frac{m_X - 1}{k_X - 1} \right)^{-1} \cdot \alpha_k^{[X]} \cdot Q_t.$$
5. *Parameter Estimation – Impact Vectors* (ref section 2.7.1 page 37). The system total impact vectors are calculated for each failure cause within the database.

6. *Parameter Estimation – Partial Alpha Factor Model* (ref section 2.7.2 page 44).

The partial alpha factors for each failure cause are calculated. The gamma parameter representing the frequency of each cause is calculated. The alpha factor for each CCCG (i.e  $\alpha_2^{[Y]}$ ) is calculated.

7. *System Quantification and Results Interpretation*. The remainder of the CCF analysis process is identical to using the AFM.

### **6.3. Parameter Description**

The Partial Alpha Factor Model has two parameter types, partial alpha factors and gamma factors.

Partial alpha factors are the alpha factors which are calculated for each coupling factor. Each partial alpha factor represents the strength of that coupling factor to transmit a failure to other components. For example  $\alpha_2$  is further split into  $\alpha_{2,CF1} \dots \alpha_{2,CFi} \dots \alpha_{2,CFw}$ , where  $CFi$  is the  $i^{th}$  coupling factor.  $0 \leq i \leq w$ .

Gamma factors, represent the portion of system failures which have the potential to propagate through the coupling factor. For example  $\gamma_{CF1}$  is the portion of failures which have the potential to propagate through Coupling Factor 1.

### **6.4. Parameter Estimation**

NUREG/CR-6823 discusses a number of methods for conducting data analysis and

parameter assessments, however this section will use two formulations of parameter estimation, a classical/frequentist interpretation using Maximum Likelihood Estimates and Bayesian methodology with conjugate priors.

### 6.4.1. Partial Alpha Factor

#### Classical Estimation

The frequentist point estimate for the partial alpha factor is:

$$\hat{\alpha}_{k,i} = \frac{n_{k,i}}{n_{p,i}}$$

$\alpha_{k,i}$  = a partial alpha factor which represents the portion of system failure events which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ) when there was a potential for failure propagation through coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$

$n_{k,i}$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ) of coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$

$n_{p,i}$  = the total number of failure events/frequency which had the opportunity for the failure to propagate through coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$ .

When the failure cause taxonomy is defined in such a way that each cause could only propagate through one coupling factor (the topic of Chapter 4), this estimate becomes:

$$\hat{\alpha}_{k,i} = \frac{n_{k,i}}{n_{t,i}}$$

where

$$n_{t,i} = \sum_{k=1}^m n_{k,i}$$

$n_{t,i}$  = *the total number of common cause failure events for coupling factor/cause  $i$  where  $i \in \{1,2,3,\dots,w\}$ .*

### Bayesian Estimations

Where a Beta distribution prior is used,  $\pi^0(\alpha_{k,i}; a_{k,i}^0, b_{k,i}^0)$ , the parameters for the posterior distribution of the partial alpha factor,  $\pi(\alpha_{k,i}; a_{k,i}, b_{k,i})$ , is:

$$a_{k,i} = a_{k,i}^0 + n_{k,i}$$

$$b_{k,i} = b_{k,i}^0 + n_{p,i} - n_{k,i}$$

When the failure cause taxonomy is defined in such a way that each cause could only propagate through one coupling factor (the topic of Chapter 4), then  $n_{p,i} = n_{t,i}$  and  $\sum_{k=1}^m \alpha_{k,i} = 1$ . Now the partial alpha factor parameter vector for each cause,  $\alpha_i$ , can be modeled with a Dirichlet distribution:

$$\alpha_i \sim \text{Dirichlet}_m\{\psi_i\}$$

$\alpha_i$  = *the portion of failure events for each multiplicity of failure  $[\alpha_{1,i}, \alpha_{2,i}, \dots, \alpha_{m,i}]$  for failure cause  $i$ .*

$\psi_i$  = *the equivalent count of failure events for each multiplicity of failure with cause  $I$   $[\psi_{1,i}, \psi_{2,i}, \dots, \psi_{m,i}]$*

The point estimates for each partial alpha factor can be obtained using:

$$\hat{\alpha}_{k,i} = \frac{\psi_{k,i}}{\sum_{k=1}^m \psi_{k,i}}$$

The parameter  $\alpha_i$ , which is the unknown of interest (UOI), can be estimated using Bayes' rule:

$$f(\alpha_i|\mathbf{n}_i) = \frac{f(\alpha_i)L(\mathbf{n}_i|\alpha_i)}{\sum_{j=1}^m L(\mathbf{n}_i|\alpha_{j,i})f(\alpha_{j,i})}$$

Where:

- $f(\alpha_i)$  = is the prior distribution of the parameter  $\alpha_i$
- $L(\mathbf{n}_i|\alpha_i)$  = is the likelihood equation for observing the evidence  $\mathbf{n}_i$  given the parameters  $\alpha_i$ .
- $f(\alpha_i|\mathbf{n}_i)$  = the posterior distribution of  $\alpha_i$  given the evidence  $\mathbf{n}_i$
- $\mathbf{n}_i$  = the number of failure events for each multiplicity of failure  $[n_{1,i}, n_{2,i}, \dots, n_{m,i}]$  for failure cause  $i$ .
- $\alpha_i$  = the portion of failure events for each multiplicity of failure  $[\alpha_{1,i}, \alpha_{2,i}, \dots, \alpha_{m,i}]$  for failure cause  $i$ .

The likelihood equation of observing the number of failures in each failure cause category,  $\mathbf{n}_i = [n_{1,i}, n_{2,i}, \dots, n_{m,i}]$  is distributed as a multinomial distribution with parameters  $\alpha_i = [\alpha_{1,i}, \alpha_{2,i}, \dots, \alpha_{m,i}]$ .

$$\mathbf{n}_i \sim \text{Multinomial}_m\{n_{t,i}, \alpha_i\}$$

where

$$n_{t,i} = \text{the number of failure events with cause } i$$

Therefore the hyper parameters,  $\psi_i$ , for the posterior  $\alpha_i$  given evidence  $\mathbf{n}_i$  using a Dirichlet prior with parameters,  $\psi_{i,0}$ , is:

$$\psi_i = \psi_{i,0} + \mathbf{n}_i$$

The choice of a prior distribution parameters,  $\psi_{i,0}$ , depends on the availability of data and will be discussed in section 6.5.

### 6.4.2. Gamma Factor

#### Classical Estimation

The frequentist point estimate for the gamma factor is:

$$\hat{\gamma}_i = \frac{n_{p,i}}{n_t}$$

$\gamma_i =$  the portion of failure events which had the potential to propagate through coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$   
 $n_t =$  the total number of failure events/frequency.

When the failure cause taxonomy is defined in such a way that each cause could only propagate through one coupling factor (the topic of Chapter 4), this estimate becomes:

$$\gamma_i = \frac{n_{t,i}}{n_t}$$

and

$$\sum_{i=1}^w \gamma_i = 1$$

#### Bayesian Estimations

Where a Beta distribution prior is used,  $\pi^0(\gamma_i; a_i^0, b_i^0)$ , the parameters for the posterior distribution of the partial alpha factor,  $\pi(\gamma_i; a_i, b_i)$ , is:

$$a_i = a_i^0 + n_{p,i}$$

$$b_i = b_i^0 + n_t - n_{p,i}$$

When the failure cause taxonomy is defined in such a way that each cause could only propagate through one coupling factor (the topic of Chapter 4), then  $n_{p,i} = n_{t,i}$  and  $\sum_{i=1}^w \gamma_i = 1$ . Now the gamma parameter vector,  $\gamma$ , can be modeled with a Dirichlet distribution:

$$\boldsymbol{\gamma} \sim \text{Dirichlet}_w\{\boldsymbol{\varphi}\}$$

$\boldsymbol{\gamma} =$  the portion of failure events for each cause  $[\gamma_1, \gamma_2, \dots, \gamma_w]$   
 $\boldsymbol{\varphi} =$  the equivalent count of failure events for each cause  
 $[\varphi_1, \varphi_2, \dots, \varphi_w]$

The point estimates for each gamma factor can be obtained using:

$$\hat{\gamma}_i = \frac{\varphi_i}{\sum_{i=1}^w \varphi_i}$$

The parameter  $\boldsymbol{\gamma}$ , which is the unknown of interest (UOI), can be estimated using

Bayes' rule:

$$f(\boldsymbol{\gamma}|\mathbf{n}_t) = \frac{f(\boldsymbol{\gamma})L(\mathbf{n}_t|\boldsymbol{\gamma})}{\sum_{j=1}^w L(\mathbf{n}_t|\boldsymbol{\gamma}_j)f(\boldsymbol{\gamma}_j)}$$

Where:

$f(\boldsymbol{\gamma}) =$  is the prior distribution of the parameter  $\boldsymbol{\gamma}$   
 $L(\mathbf{n}_t|\boldsymbol{\gamma}) =$  is the likelihood equation for observing the evidence  
 $f(\boldsymbol{\gamma}|\mathbf{n}_t) =$  the posterior distribution of  $\boldsymbol{\gamma}$  given the evidence  $\mathbf{n}_t$   
 $\mathbf{n}_t =$  the number of failure events for each cause  $[n_{t,1}, n_{t,2}, \dots, n_{t,w}]$   
 $\boldsymbol{\gamma} =$  the portion of failure events for each cause  $[\gamma_1, \gamma_2, \dots, \gamma_w]$

The likelihood equation of observing the number of failures in each failure cause category,  $\mathbf{n}_t = [n_{t,1}, n_{t,2}, \dots, n_{t,w}]$  is distributed as a multinomial distribution with parameters  $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_w]$ .

$$\mathbf{n}_t \sim \text{Multinomial}_w\{n_t, \boldsymbol{\gamma}\}$$

where

$n_t =$  the number of failure events

Therefore the hyper parameters,  $\boldsymbol{\varphi}$ , for the posterior  $\boldsymbol{\gamma}$  given evidence  $\mathbf{n}_t$  using a

Dirichlet prior with parameters,  $\boldsymbol{\varphi}_0$ , is:

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}_0 + \mathbf{n}_t$$

The choice of a prior distribution parameters,  $\boldsymbol{\varphi}_0$ , depends on the availability of data and will be discussed in section 6.5.

### 6.4.3. Assessed Alpha Factor

Partial alpha factors may be converted back to alpha factors for use with CCBEs from each CCCG. This is done through summing the contributions from each coupling factor.

$$\alpha'_k = \sum_{i \in r} \gamma_i \alpha_{k,i}$$

$r$  = the coupling factors shared by the components within the CCCG being analysed,  $r \subseteq \{1, 2, 3, \dots, w\}$ .

$\alpha'_k$  = the assessed alpha factor. This is the system alpha factor which only considers the coupling factors shared by the components within the CCCG where  $2 \leq k \leq m$

$\alpha_1$  is the single failures and any contribution from coupling factors which are not shared.

$$\alpha'_1 = 1 - \sum_{CCCGs} \sum_{k=2}^m \alpha'_k$$

When components share all coupling factors (complete symmetry) this alpha factor estimate is equivalent to the AFM. Where components only share some coupling factors, this estimate of the alpha factor is reduced commensurate with the benefits of decoupling that particular feature. This will be demonstrated through the use of examples in section 6.6.



## 6.5. Parameter Quantification

In estimating the PAFM parameters from data, it is assumed that the failure taxonomy allows for a one to one direct relationship between failure causes and coupling factors. This issue is discussed in Chapter 4. Occasions where a failure cause may propagate over multiple coupling factors are discussed in section 6.8.1.

### 6.5.1. *Using Impact Vectors*

Where component specific data exists, the evidence required to calculate the parameter estimations can be quantified using the impact vector methodology.

The average impact vector for a CCF event can be represented with the inclusion of the failure cause:

$$\bar{I} = [\bar{F}_0, \bar{F}_1, \dots, \bar{F}_m][\text{Cause}]$$

Then the sum of average impact vectors for J events for a particular cause is:

$$n_{\Omega,i} = [n_{1,i}, \dots, n_{m,i}]$$

where

$$n_{k,i} = \sum_{j=1}^J \bar{F}_k(j)[\text{Cause}]$$

*where Cause=i*

$n_{k,i}$  = the total number of CCF basic events caused by  $i$  involving the failure of  $k$  similar components.

Note,  $n_0$  is not included because a failure cause cannot be determined when there was no failure.

The quantities required to estimate the partial alpha factors and gamma factors can now be calculated as:

$$n_{t,i} = \sum_{k=1}^m n_{k,i}$$

$$n_t = \sum_{i=1}^m n_{t,i} = \sum_{k=1}^m n_k$$

$$\mathbf{n}_t = [n_{t,1}, n_{t,2}, \dots, n_{t,w}]$$

$$\mathbf{n}_i = [n_{1,i}, n_{2,i}, \dots, n_{m,i}]$$

Using impact vectors to quantify the PAFM allows for the ability to use currently accepted mapping rules from systems with different CCCG sizes, include uncertainty around observed partial failures, coupling factors and time delays as provided in the impact vector methodology.

### ***6.5.2. Using Generic Data Sources***

Where detailed information about each failure cause is unknown the analyst may use data from generic data sources or from similar systems to create a prior distribution before incorporating system specific data. In doing so the analyst must be careful not to overstate the strength of the generic data to estimate the specific application.

For example, a generator has been installed in a two train configuration, as per example

1; however the installation design is not common across industry. Both generators have been running continuously for 250 days (500 days in total). During that time there have been two failures, one caused by the installation procedure, and the other due to the external environment.

Plant specific data is:

$$n_{\Omega,IP} = [0, 1, 0]$$

$$n_{\Omega,MH} = [0, 0, 0]$$

$$n_{\Omega,EE} = [0, 1, 0]$$

$$n_0 = 498$$

The point estimate using the plant specific data is:

$$Q_t = \frac{n_F}{N_1} = \frac{2}{500} = 0.004$$

$$\alpha_{2,IP} = \frac{0}{1+0} = 0$$

$$\gamma_{IP} = \frac{1}{1+0+1} = 0.5$$

$$\alpha_{2,MH} = \frac{0}{0+0} = 0$$

$$\gamma_{MH} = \frac{0}{1+0+1} = 0$$

$$\alpha_{2,EE} = \frac{0}{1+0} = 0$$

$$\gamma_{EE} = \frac{1}{1+0+1} = 0.5$$

The estimates from the plant specific data show a lack of data to estimate partial alpha factors and gamma factors. The analyst consults the CCFDB for generator data, as detailed in Table 7 (example 1 EDG data). This data shows a total of 35 failures in 5834 days of operation. The generic data set is more plentiful than the plant specific data,

however the analyst believes the generic data is not completely representative due to this generator's unique configuration.

A summary of the generic data is:

$$n_{\Omega,IP}^{[E]} = [0, 3, 1]$$

$$n_{\Omega,MH}^{[E]} = [0, 14, 2]$$

$$n_{\Omega,EE}^{[E]} = [0, 8, 2]$$

$$n_0^{[E]} = 2887$$

The point estimate using the generic data is:

$$Q_t = \frac{n_F^{[E]}}{N_1^{[E]}} = \frac{35}{5834} = 0.006$$

$$\alpha_{2,IP}^{[E]} = \frac{1}{3 + 1} = 0.25$$

$$\gamma_{IP}^{[E]} = \frac{4}{4 + 16 + 10} = 0.13$$

$$\alpha_{2,MH}^{[E]} = \frac{2}{14 + 2} = 0.125$$

$$\gamma_{MH}^{[E]} = \frac{16}{4 + 16 + 10} = 0.53$$

$$\alpha_{2,EE}^{[E]} = \frac{2}{8 + 2} = 0.2$$

$$\gamma_{EE}^{[E]} = \frac{10}{4 + 16 + 10} = 0.33$$

The data from both information sources can be combined through Bayesian updating. However, in order for the plentiful generic data not to dominate the estimate, it should be adjusted by a weighting factor  $w_e$  which is an engineering assessment of the strength between the component similarities and operating context for generic data against the target system.

Therefore the hyper-parameters of the posterior distribution for the partial alpha factors can be calculated as:

$$\boldsymbol{\psi}_i = \boldsymbol{\psi}_{i,0} + \sum_{e=1}^E w_e \mathbf{n}_{i,e}$$

$\boldsymbol{\psi}_i =$  the equivalent count of failure events for each multiplicity of failure with cause  $I$  [ $\psi_{1,i}, \psi_{2,i}, \dots, \psi_{m,i}$ ]

$\mathbf{n}_{i,e} =$  the number of failure events for each multiplicity of failure [ $n_{1,i}, n_{2,i}, \dots, n_{m,i}$ ] for failure cause  $i$  from evidence source  $e$ .

$w_e =$  evidence weighting factor

Therefore the hyper-parameters of the posterior distribution for the gamma factors can be calculated as:

$$\boldsymbol{\varphi} = \boldsymbol{\varphi}_0 + \sum_{e=1}^E w_e \mathbf{n}_{t,e}$$

$\boldsymbol{\varphi} =$  *the equivalent count of failure events for each cause*

$[\varphi_1, \varphi_2, \dots, \varphi_w]$

$\mathbf{n}_{t,e} =$  *the number of failure events for each cause  $[n_{t,1}, n_{t,2}, \dots, n_{t,w}]$  from source  $e$ .*

$w_e =$  *evidence weighting factor*

The analyst makes an engineering assessment that  $w_e$  for the generic data is 0.3. This avoids non-zero elements, allows for the generic data to influence the estimate but not dominate the plant specific data. The following example shows the calculation for the EE alpha factor using a Novick and Hall improper prior (see section 6.5.4):

$$\boldsymbol{\psi}_{EE,0} = [0, 0]$$

$$n_{\Omega,EE}^{[E]} = [8, 2], \quad w_{e1} = 0.3$$

$$n_{\Omega,EE} = [1, 0], \quad w_{e2} = 1$$

Now:

$$\begin{aligned} \boldsymbol{\psi}_i &= \boldsymbol{\psi}_{i,0} + \sum_{e=1}^E w_e \mathbf{n}_{i,e} \\ &= [0, 0] + \left[ \frac{12}{5}, \frac{3}{5} \right] + [1, 0] \\ &= [3.4, 0.6] \end{aligned}$$

The point estimate for the partial alpha factor is now:

$$\alpha_{2,EE} = \frac{0.6}{3.4 + 0.6} = 0.15$$

The following example shows the calculation for the EE gamma factor using a Novick and Hall improper prior (see section 6.5.4):

$$\boldsymbol{\varphi}_0 = [0, 0, 0]$$

$$\mathbf{n}_{t,e2}^{[E]} = [4, 16, 10], \quad w_{e1} = 0.3$$

$$\mathbf{n}_{t,e2} = [1, 0, 1], \quad w_{e2} = 1$$

Now:

$$\begin{aligned} \boldsymbol{\varphi} &= \boldsymbol{\varphi}_0 + \sum_{e=1}^E w_e \mathbf{n}_{t,e} \\ &= [0, 0, 0] + \left[ \frac{6}{5}, \frac{24}{5}, 3 \right] + [1, 0, 1] \\ &= [2.2, 5.8, 4] \end{aligned}$$

The point estimate for the EE gamma factor is now:

$$\hat{\gamma}_{EE} = \frac{4}{2.2 + 5.8 + 4} = 0.33$$

The engineering assessment of the weighting factor affects the strength of the data from a particular evidence source. Table 19 shows the effect which different weighting factors have in the estimation of  $\alpha_{2,EE}$  and  $\gamma_{EE}$ . A strong weighting factor has the posterior point estimate more towards the estimate from the generic data source, a weak

prior allows for non-zero estimates which allow the plant specific data to dominate.

**Table 19: Comparison of weighting factor strengths**

$\alpha_{2,EE}$ Plant Specific	$\alpha_{2,EE}$ Generic Data	$w_{e1}$ Weight Factor	Posterior $\alpha_{2,EE}$ Point Estimate	$\gamma_{EE}$ Plant Specific	$\gamma_{EE}$ Generic Data	$w_{e1}$ Weight Factor	Posterior $\gamma_{EE}$ Point Estimate
0	0.2	0	0.00	0.5	0.33	0	0.5
0	0.2	0.2	0.13	0.5	0.33	0.2	0.375
0	0.2	0.4	0.16	0.5	0.33	0.4	0.35714 3
0	0.2	0.6	0.17	0.5	0.33	0.6	0.35
0	0.2	0.8	0.18	0.5	0.33	0.8	0.34615 4
0	0.2	1	0.18	0.5	0.33	1	0.34375

### 6.5.3. Informative Prior Distributions

The objective of the prior distribution is to capture the belief the analyst has about the Unknown Of Interest (UOI). Where the prior distribution captures an assessment of the analyst's believe, it is known as a subjective prior. A detailed treatment of informative priors used in PRA analysis is provided in (Siu & Kelly 1998).

### Population Variability Prior

A common method for formulating an informative prior is to construct a distribution of the variability of the UOI across the population. Methods such as the Two-stage Bayes, Hierarchical Bayes and Empirical Bayes methods achieve this with subtle differences in their approach (Siu & Kelly 1998). The procedure is to estimate the variability of the UOI across a population. For example the PAFM parameters may be



calculated for each plant within the USA. A histogram of the plan values is then created and a distribution fitted. This then forms the prior distribution, and the collection of the plant specific data has the UOI estimate converge towards where that specific plant fits into the population variability function. The method of dividing the population into sub-populations could be applied across different component types, plant types, plant locations, component manufacturers etc.

This approach has the advantage of informing the estimate of parameters based on the variability seen across the population. For example alpha factors are generally between 0.2 and 0.001 and using this method would provide a quantitative assessment of this range.

The method for calculating the population variability function which becomes the prior requires industry wide data collection, such as the CCFDB. The specific method for dividing the data would require analysis to minimize variability.

### **Expert Elicitation**

The procedure for elicitation of quantitative estimates from experts has been proposed by multiple authors (**Cooke 1991; O'Hagan et al. 2006; Zitrou 2006b**). In addition many mathematical techniques have been proposed to combine expert's estimates, treat bias, and account for overlapping knowledge (Garthwaite et al. 2005; Kadane & Wolfson 1998; Mumpower & Stewart 1996; Skjong & Wentworth 2001) The NRC has

provided practical guidance on expert elicitation through (Meyer & Booker 1990; Ronald et al. 2005)

Two important characteristics must be quantified when using an expert's assessment of PAFM values.

- *Cause Frequency.* The portion of failures which occur due to a specific cause is affected by the gamma parameter. Due to the operating, environmental and design context of each system the likelihood of some causes may approach zero. For example items which do not undergo maintenance will have a zero probability of a maintenance procedural or maintenance human failure cause.
- *Strength of Failure Propagation.* The partial alpha factor represents the likelihood that a failure will propagate to other components. The deterministic relationship of coupling factors is captured in the qualitative assessment during the analysis procedure. However where coupling factors exist, system specific characteristics can be captured through the expert elicitation process.

#### **6.5.4. *Non-informative Prior Distributions***

When prior knowledge is vague, it is often desirable to have a repeatable, defensible prior which represents our lack of knowledge of the system. (Atwood 1996) It is particularly easy to justify a non-informative prior where there is a substantial amount of information which will make the assumptions used to form a prior insignificant.

However, when limited evidence is available, the prior can dominate the estimate and can be difficult to justify. (Siu & Kelly 1998) A detailed account for the philosophy, purpose and construction of non-informative priors can be found in (Kass & Wasserman 1996) and use of non-informative priors in PRA are described in (Siu & Kelly 1998). This section will briefly describe the non-informative Dirichlet priors that can be used for the estimation of the partial alpha factors and gamma factors.

Despite these non-informative priors allowing a repeatable prior distribution, the analyst must still apply engineering judgment to ensure that where causes cannot exist within the specific system, the gamma factor for that cause remains zero. The occasion when failure propagation cannot occur (alpha factors zero) can be accounted for in the creation of CCCGs.

### **Uniform Distribution**

The uniform distribution assigns an equal probability to each unknown of interest (UOI). This is known as the principle of indifference. and can be represented with the Dirichlet parameters,  $\theta_0$  equal to (Yang & Berger 1998):

$$\pi_0(\theta) = \text{Dirichlet}_d([1,1, \dots, 1])$$

### **Jeffreys Prior Distribution**

Proposed by Jeffrey in 1961, this prior is calculated as  $\pi_0(\theta) = \sqrt{\det(\mathbf{I}_\theta)}$  where  $\mathbf{I}_\theta$  is the Fisher information matrix. This derivation is motivated by the fact that it is not dependent upon a set of parameter variables that is chosen to describe the parameter space. The Jeffrey Prior parameters for the Dirichlet distribution are:

$$\pi_0(\theta) = \text{Dirichlet}_d\left(\left[\frac{1}{2}, \frac{1}{2}, \dots, \frac{1}{2}\right]\right)$$

### **Novick and Hall's Improper Prior**

Novick and Hall defined an “indifference prior” by selecting the parameters of a conjugate prior which is improper and provides the minimum necessary samples to the posterior (Kass & Wasserman 1996). An improper prior distribution is where the distribution does not integrate to 1. The parameters for the Dirichlet which meet Novich and Hall's criteria is:

$$\pi_0(\theta) = \text{Dirichlet}_d([0,0, \dots, 0])$$

The analyst must be careful when using such a prior as it can lead to zero estimates of the UOI,  $\theta$ .

#### ***6.5.5. Using Alpha Factors as a Constraint***

This chapter so far has discussed methods to quantify PAFM parameters where very little or no data exists. There may however be occasions where the alpha factors are

known, but the failure causes for the impact vectors is unknown. In this unique situation the same procure of using priors, expert opinions, generic data sources and plan specific data sources may be used, with the additional constraint that:

$$\alpha''_k = \sum_i \gamma_i \alpha_{k,i}$$

where

$\alpha''_k =$  *the known alpha factor for a specific system*

There are two primary methods through which adjustments can be made for ensuring this constraint is met:

- Adjust the gamma factor distribution
- Adjust the partial alpha factor distribution

### **Adjust the Gamma Factor Distribution**

The most likely reason for differences between generic data and the known alpha factor is due to different failure rates for each cause based on plant specific characteristics. For example plants which operate in different temperature ranges or use different water sources will have different distributions for each failure cause. Using an engineering assessment about the specific differences of a particular system, the analyst can adjust the gamma factors such that the alpha factor constraint is met.

## Adjust the Partial Alpha Factors

The partial alpha factors represent the strength of failure propagation given the opportunity. This factor can change between specific systems based on the soft defenses of the system such as degrees of separation, or differences between components. Using an engineering assessment about the specific differences of a particular system, the analyst can adjust the gamma factors such that the alpha factor constrain is met.

### 6.6. PAFM in System Analysis

This section describes the system analysis procedure when using the PAFM. The two examples used to demonstrate this procedure are the same two examples described in Chapter 2. A two train EDG system with perfect symmetry and a system with two EDG and three pumps with mixed redundancy.

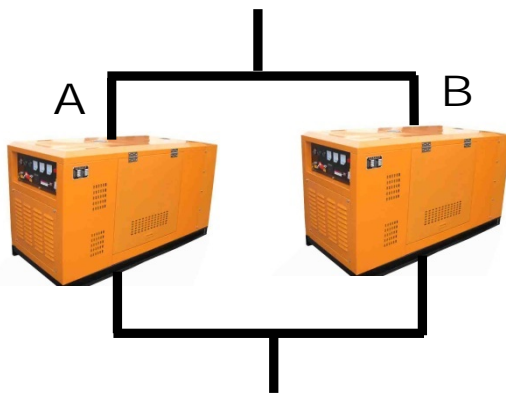


Figure 27: Reliability block diagram for example 1- Two train EDG system

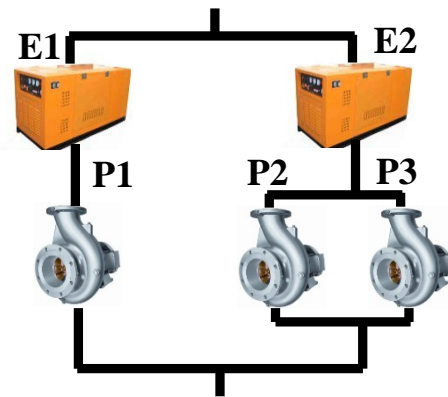


Figure 28: RBD for example 2, two EDG three pump system.

This section focuses on calculating the system probability and may skip some qualitative steps of the process for simplicity.

### 6.6.1. Qualitative Analysis

The purpose of the qualitative analysis is to form common cause component groups (CCCGs) which group components together based on their shared coupling factors.

A component may be a part of multiple CCCGs based on one or more shared attributes. However a component cannot be part of multiple CCCGs for the one attribute. Components which share a CCCG must have the same failure rate/probability. This results in the PAFM being able to model asymmetrical coupling factors but not asymmetrical components and failure rates/probabilities.

The qualitative assessment for example 1 and example 2 are in Table 20 and Table 21 respectively.

**Table 20: Qualitative dependency assessment for example 1**

<b>Component</b>	<b>Install Procedure</b>	<b>Maintenance Staff</b>	<b>Location</b>
EDG 1 (A)	EDG IP	Team X	Room Y
EDG 2 (B)	EDG IP	Team X	Room Y

**Table 21: Qualitative dependency assessment for example 2**

<b>Component</b>	<b>Install Procedure</b>	<b>Maintenance Staff</b>	<b>Location</b>
------------------	--------------------------	--------------------------	-----------------

EDG 1 (E1)	EDG	Team X	Room Y
EDG 2 (E2)	EDG	Team X	Room Y
Pump 1 (P1)	Pump V1.1	Team X	Room Y
Pump 2 (P2)	Pump V2.8	Team X	Room Y
Pump 3 (P3)	Pump V1.1	Team Y	Room X

In example 1 it is clear that the two EDG share all coupling factors and should form one CCCG.

$$CCCG^{[E]} = \{A, B\}$$

In example 2 while the EDGs are again symmetrical, pumps 1 and 2 only share the same maintenance team and location, pumps 1 and 3 share an installation procedure. Therefore pump 1 will become part of two CCCGs. The EDG and pumps do share features, however due to the inability to model asymmetrical component failure rates, these dependencies cannot be modeled (as per the AFM). Therefore the CCCGs for example 2 are:

$$CCCG_X = \{E_1, E_2\}$$

$$CCCG_Z = \{P_1, P_2\}$$

$$CCCG_Y = \{P_1, P_3\}$$

### ***6.6.2. Identification of Common Cause Basic Events***

The CCBEs will be constructed with consideration for all CCCGs which the component is a part of. If a component was part of multiple CCCGs, then the CCBEs for both CCCGs would be added.



The CCBE events for example 1 and example 2 are shown in Table 22 and Table 23 respectively.

**Table 22: CCBE for example 1**

<b>Component</b>	<b>Common Cause Basic Events</b>
EDG 1 (A)	$A_i, X_{AB}$
EDG 2 (B)	$B_i, X_{AB}$

**Table 23: CCBE for example 2**

<b>Component</b>	<b>Common Cause Basic Events</b>
EDG 1 ( $E_1$ )	$E_{1,i}, X_{E_1,E_2}$
EDG 2 ( $E_2$ )	$E_{2,i}, X_{E_1,E_2}$
Pump 1 ( $P_1$ )	$P_{1,i}, Z_{P_1,P_2}, Y_{P_1,P_3}$
Pump 2 ( $P_2$ )	$P_{2,i}, Z_{P_1,P_2}$
Pump 3 ( $P_3$ )	$P_{3,i}, Y_{P_1,P_3}$

### **6.6.3. Incorporate into Fault Tree**

The CCBEs are incorporated into the fault tree as per Chapter 2.

The fault tree for example 1 after substitution of CCBEs is shown in Figure 29.

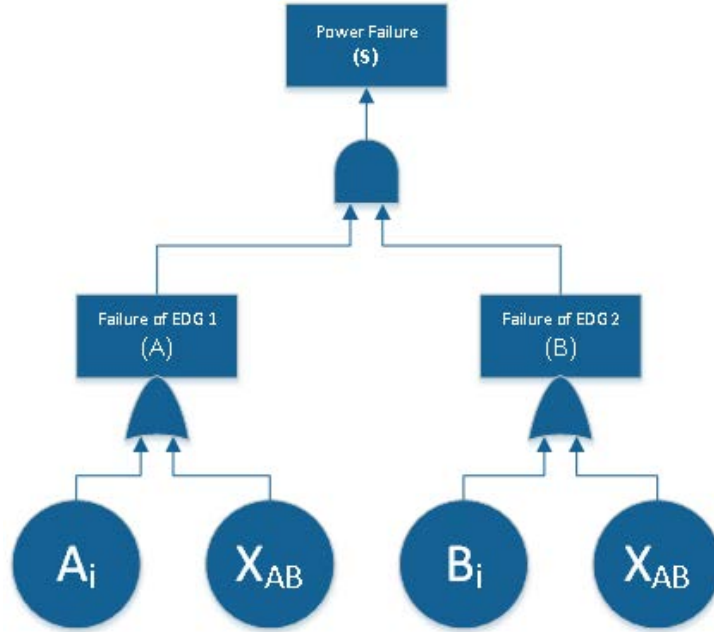


Figure 29: Fault tree for example 1 with CCBEs

The cut sets for example 1 are now:

$$\{A_i, B_i\}; \{X_{AB}\}$$

The fault tree for example 2 after substitution of CCBEs is shown in Figure 30.

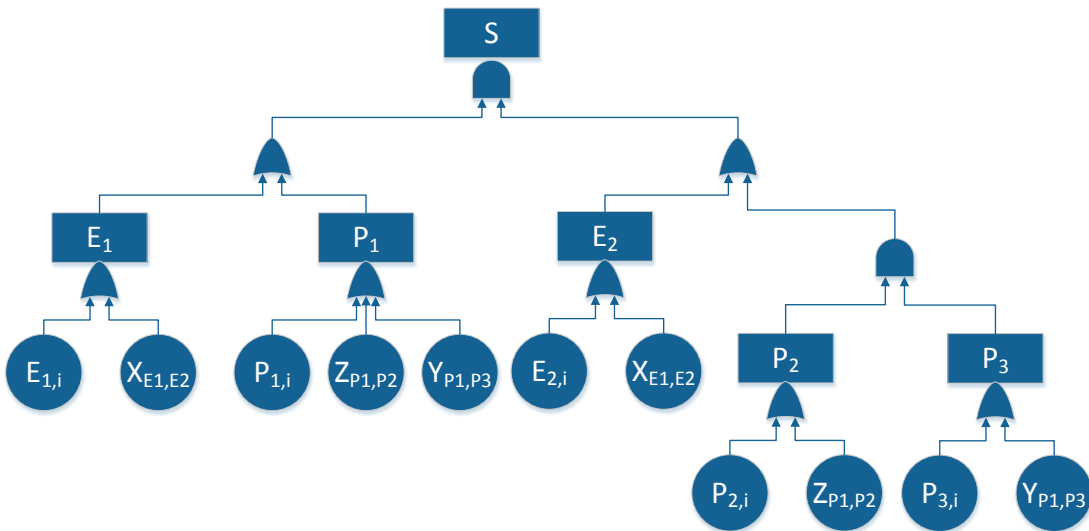


Figure 30: Fault tree for example 2 with CCBEs

The cut sets for example 2 with CCBEs are:

$$\begin{aligned} & \{E_{1,i}, E_{2,i}\}; \{P_{1,i}, E_{2,i}\}; \{P_{1,i}, P_{2,i}, P_{3,i}\}; \{E_{1,i}, P_{2,i}, P_{3,i}\}; \\ & \{P_{3,i}, Z_{P1,P2}\}; \{E_{2,i}, Z_{P1,P2}\}; \{P_{2,i}, Y_{P1,P3}\}; \{E_{2,i}, Y_{P1,P3}\}; \\ & \{Z_{P1,P2}, Y_{P1,P3}\}; \{X_{E1,E2}\} \end{aligned}$$

Note that the cut set  $\{Z_{P1,P2}, Y_{P1,P3}\}$  treats the failure of a single component due to different causes in separate CCCGs as independent, not mutually exclusive. This is the same assumption used to separate individual and common cause basic events and may have minimal influence.

One problem with this approach is that the CCBEs are treated as independent events, instead of mutually exclusive. This

#### **6.6.4. Parametric representation of CCBEs**

The parameter representation of the CCBEs are the same as Chapter 2. For example 1 the CCBEs are equal to:

$$P(A_i) = P(B_i) = Q_1^{(2)}$$

$$P(X_{AB}) = Q_2^{(2)}$$

The example 1 system equation is now:

$$P(S) = P(A_i)P(B_i) + P(X_{AB})$$

$$P(S) = \left(Q_1^{(2)}\right)^2 + Q_2^{(2)}$$

For example 2 the CCBEs are equal to:

$$P(E_{1,i}) = P(E_{2,i}) = Q_1^{(2)[E]}$$

$$P(X_{E1,E2}) = Q_2^{(2)[X]}$$

$$P(P_{1,i}) = Q_1^{(2)[P_1]}$$

$$P(P_{2,i}) = Q_1^{(2)[P_2]}$$

$$P(P_{3,i}) = Q_1^{(2)[P_3]}$$

$$P(Z_{P1,P2}) = Q_2^{(2)[Z]}$$

$$P(Y_{P1,P3}) = Q_2^{(2)[Y]}$$

In comparison to the AFM, the pumps are not symmetrical and so:

$$Q_1^{(2)[P_1]} \neq Q_1^{(2)[P_2]} \neq Q_1^{(2)[P_3]}$$

Using rare event approximation, the example 2 system equation is now:

$$\begin{aligned} P(S) = & \left( Q_1^{(2)[E]} \right)^2 + Q_1^{(2)[P_1]} Q_1^{(2)[E]} + Q_1^{(2)[P_1]} Q_1^{(2)[P_2]} Q_1^{(2)[P_3]} + Q_1^{(2)[E]} Q_1^{(2)[P_2]} Q_1^{(2)[P_3]} \\ & + Q_1^{(2)[P_3]} Q_2^{(2)[Z]} + Q_1^{(2)[E]} Q_2^{(2)[Z]} + Q_1^{(2)[P_2]} Q_2^{(2)[Y]} + Q_1^{(2)[E]} Q_2^{(2)[Y]} \\ & + Q_2^{(2)[Z]} Q_2^{(2)[Y]} + Q_2^{(2)[X]} \end{aligned}$$

### 6.6.5. *Partial Alpha Factor Model Parameterization*

The Basic Parameter values are quantified using assessed alpha factors which only add the contribution from the relevant dependencies.

First, the CCBEs are quantified as per Chapter 2, except using assessed alpha factors.

$$Q_k^{(m)} = \binom{m-1}{k-1}^{-1} \cdot \alpha'_k \cdot Q_t$$

For example 1 the basic events for each component are:

$$Q_1^{(2)} = \alpha'_1 Q_t, \quad Q_2^{(2)} = \alpha'_2 \cdot Q_t$$

Due to the components being in perfect symmetry, the assessed alpha factors have a contribution from all coupling factors.

$$\alpha'_2 = \sum_{i=\{IP,MH,EE\}} \gamma_i \alpha_{2,i}$$

$$\alpha'_1 = 1 - \alpha'_2$$

For example 2 the basic events for each component are:

$$Q_1^{(2)[P]} = \alpha_1^{[P]} Q_t^{[P]}, \quad Q_2^{(2)[P]} = \alpha_2^{[P]} \cdot Q_t^{[P]}$$

For example 2 the EDGs are symmetrical and therefore the basic events are:

$$Q_1^{(2)[E]} = \alpha_1^{[E]} Q_t^{[E]}, \quad Q_2^{(2)[X]} = \alpha_2^{[X]} \cdot Q_t^{[E]}$$

$$\alpha_2^{[X]} = \sum_{i=\{IP,MH,EE\}} \gamma_i^{[E]} \alpha_{2,i}^{[E]}$$

$$\alpha_1^{[E]} = 1 - \alpha_2^{[X]}$$

For example 2, the pumps in  $CCCG_Z$  only share the maintenance team and external environment. The assessed alpha factors will not include the contribution from installation procedures.

$$Q_2^{(2)[Z]} = \alpha_2^{[Z]} \cdot Q_t^{[P]}$$

$$\alpha_2^{[Z]} = \sum_{i=\{MH,EE\}} \gamma_i^{[P]} \alpha_{2,i}^{[P]}$$

The pumps in  $CCCG_Y$  only share installation procedure and the assessed alpha factor will not include the contribution from maintenance team and external environment.

$$Q_2^{(2)[Y]} = \alpha_2'^{[Y]} \cdot Q_t^{[P]}$$

$$\alpha_2'^{[Y]} = \sum_{i=\{IP\}} \gamma_i^{[P]} \alpha_{2,i}^{[P]}$$

The assessed alpha factors for each pump may now be calculated as the remaining failure events.

$$Q_1^{(2)[P_1]} = \alpha_1'^{[P_1]} \cdot Q_t^{[P]}, \quad \alpha_1'^{[P_1]} = 1 - \alpha_2'^{[Z]} - \alpha_2'^{[Y]}$$

$$Q_1^{(2)[P_2]} = \alpha_1'^{[P_2]} \cdot Q_t^{[P]}, \quad \alpha_1'^{[P_2]} = 1 - \alpha_2'^{[Z]}$$

$$Q_1^{(2)[P_3]} = \alpha_1'^{[P_3]} \cdot Q_t^{[P]}, \quad \alpha_1'^{[P_3]} = 1 - \alpha_2'^{[Y]}$$

#### 6.6.6. Parameter Estimation – Impact Vectors

Impact vectors are calculated for each CCF event as per Chapter 2. The sum of the average impact vectors must be conducted for each cause.

For the examples, the data and average impact vectors for an EDG are contained in Table 7 on page 42. The sum of the EDG average impact vectors for each cause is:

$$n_{\Omega,IP}^{[E]} = [0, \quad 172.2, \quad 2.8]$$

$$n_{\Omega,MH}^{[E]} = [0, \quad 154.35, \quad 3.15]$$

$$n_{\Omega,EE}^{[E]} = [0, \quad 16.45, \quad 1.05]$$

$$n_0^{[E]} = 29400$$

The data and average impact vectors for a pump are contained in

Table 9 on page 43.

The sum of the pump average impact vectors for each cause is:

$$n_{\Omega,IP}^{[P]} = [0, \quad 26.0663, \quad 0.1838]$$

$$n_{\Omega,MH}^{[P]} = [0, \quad 59.4125, \quad 1.838]$$

$$n_{\Omega,EE}^{[P]} = [0, \quad 82.5213, \quad 4.9788]$$

$$n_0^{[P]} = 44433$$

### 6.6.7. Parameter Estimation – Partial Alpha Factor Model

In order to calculate the assessed alpha factors, the partial alpha factors and gamma factors must be calculated:

For the EDG example, the partial alpha factors and gamma factors can be calculated as:

$$\alpha_{2,IP}^{[E]} = \frac{2.8}{172.2 + 2.8} = 0.016 \qquad \gamma_{IP}^{[E]} = \frac{175}{175 + 157.5 + 17.5} = 0.5$$

$$\alpha_{2,MH}^{[E]} = \frac{3.15}{154.35 + 3.15} = 0.020 \qquad \gamma_{MH}^{[E]} = \frac{157.5}{175 + 157.5 + 17.5} = 0.45$$

$$\alpha_{2,EE}^{[E]} = \frac{1.05}{16.45 + 1.05} = 0.060 \qquad \gamma_{EE}^{[E]} = \frac{17.5}{175 + 157.5 + 17.5} = 0.05$$

The assessed alpha factors for example 1 can now be calculated as:

$$\alpha'_2 = \sum_{i=\{IP,MH,EE\}} \gamma_i \alpha_{2,i}$$

$$\begin{aligned}
&= (0.5)(0.016) + (0.45)(0.02) + (0.05)(0.06) \\
&= 0.02
\end{aligned}$$

and

$$\alpha'_1 = 1 - \alpha'_2 = 0.98$$

For the pump example, the partial alpha factors and gamma factors can be calculated

as:

$$\begin{aligned}
\alpha_{2,IP}^{[P]} &= \frac{0.18375}{26.06625 + 0.18375} = 0.007 & \gamma_{IP}^{[P]} &= \frac{26.25}{26.25 + 61.25 + 87.5} = 0.15 \\
\alpha_{2,MH}^{[P]} &= \frac{1.8375}{59.4125 + 1.8375} = 0.03 & \gamma_{MH}^{[P]} &= \frac{61.25}{26.25 + 61.25 + 87.5} = 0.35 \\
\alpha_{2,EE}^{[P]} &= \frac{4.97875}{82.52125 + 4.97875} = 0.0569 & \gamma_{EE}^{[P]} &= \frac{87.5}{26.25 + 61.25 + 87.5} = 0.50
\end{aligned}$$

The assessed alpha factors can now be calculated as:

$$\alpha'_2^{[X]} = \sum_{i=\{IP,MH,EE\}} \gamma_i^{[E]} \alpha_{2,i}^{[E]} = 0.02$$

$$\alpha'_2^{[Z]} = \sum_{i=\{MH,EE\}} \gamma_i^{[P]} \alpha_{2,i}^{[P]} = 0.03895$$

$$\alpha'_2^{[Y]} = \sum_{i=\{IP\}} \gamma_i^{[P]} \alpha_{2,i}^{[P]} = 0.00105$$

$$\alpha'_1^{[P_1]} = 1 - \alpha'_2^{[Z]} - \alpha'_2^{[Y]} = 0.96$$

$$\alpha'_1^{[P_2]} = 1 - \alpha'_2^{[Z]} = 0.96105$$

$$\alpha'_1^{[P_3]} = 1 - \alpha'_2^{[Y]} = 0.99895$$

### 6.6.8. System Quantification and Results Interpretation.



The parameter estimates may now be substituted back into the system equations.

For example 1, the system equation is:

$$P(S) = \left(Q_1^{(2)}\right)^2 + Q_2^{(2)}$$

Where:

$$Q_1^{(2)} = \alpha'_1 Q_t, \quad Q_2^{(2)} = \alpha'_2 \cdot Q_t, \quad Q_t = 0.006$$

$$\alpha'_1 = 0.98, \quad \alpha'_2 = 0.02$$

Therefore the system failure probability is:

$$\begin{aligned} P(S) &= (\alpha'_1 Q_t)^2 + \alpha'_2 \cdot Q_t \\ &= (0.98 \times 0.006)^2 + (0.02)(0.006) \\ &= 1.546e-4 \end{aligned}$$

The system failure probability for example 1 using the AFM method was 1.025e-3, which is equal to the PAFM system failure probability. Example 1 demonstrates that where each component shares all coupling factors, the PAFM and AFM results are equal.

For example 2, instead of calculating the system probability of failure using the system equation, the following quantities can be placed into the fault tree for calculation:

$$P(E_{1,i}) = P(E_{2,i}) = Q_1^{(2)[E]} = \alpha_1^{[E]} Q_t^{[E]} = 5.88e-3$$

$$P(X_{E1,E2}) = Q_2^{(2)[X]} = \alpha_2^{[X]} \cdot Q_t^{[E]} = 1.2e-4$$

$$P(P_{1,i}) = Q_1^{(2)[P1]} = \alpha_1^{[P1]} \cdot Q_t^{[P]} = 1.9584e-3$$

$$P(P_{2,i}) = Q_1^{(2)[P_2]} = \alpha_1^{[P_2]} \cdot Q_t^{[P]} = 1.9605e-3$$

$$P(P_{3,i}) = Q_1^{(2)[P_3]} = \alpha_1^{[P_3]} \cdot Q_t^{[P]} = 2.0379e-3$$

$$P(Z_{P_1,P_2}) = Q_2^{(2)[Z]} = \alpha_2^{[Z]} \cdot Q_t^{[P]} = 7.95e-5$$

$$P(Y_{P_1,P_3}) = Q_2^{(2)[Y]} = \alpha_2^{[Y]} \cdot Q_t^{[P]} = 2.1e-6$$

Substitution back into the system equation gives:

$$\begin{aligned} P(S) &= \left(Q_1^{(2)[E]}\right)^2 + Q_1^{(2)[P_1]}Q_1^{(2)[E]} + Q_1^{(2)[P_1]}Q_1^{(2)[P_2]}Q_1^{(2)[P_3]} + Q_1^{(2)[E]}Q_1^{(2)[P_2]}Q_1^{(2)[P_3]} \\ &\quad + Q_1^{(2)[P_3]}Q_2^{(2)[Z]} + Q_1^{(2)[E]}Q_2^{(2)[Z]} + Q_1^{(2)[P_2]}Q_2^{(2)[Y]} + Q_1^{(2)[E]}Q_2^{(2)[Y]} \\ &\quad + Q_2^{(2)[Z]}Q_2^{(2)[Y]} + Q_2^{(2)[X]} \\ &= 1.668e-4 \end{aligned}$$

The system failure probability/rate is calculated as 1.668e-4. This is a similar result to Chapter 2 for the AFM. The PAFM mixes assuming independence between failure causes (between CCCGs) and assuming mutually exclusive events (by addition parameters in an assessed alpha factor). This means the PAFM will provide system reliability estimates between the AFM and the GDM estimate.

### ***6.7. PAFM in Event Assessment***

While the Partial Alpha Factor Model also allows for the modeling of components which share different dependencies with different components, the primary means for calculating partial alpha factors is for use in event assessment.

The two event assessment scenarios will be presented:

- Event assessment with knowledge of a component failure
- Event assessment with knowledge of a component failure and failure cause

In order to demonstrate each method, example 1 of symmetrical EDGs will be used.

### 6.7.1. Knowledge of Failure

As the PAFM is based on the AFM, the same procedure described in Section 2.11 may be used.

The system failure probability for example 1 is given as:

$$P(S) = (\alpha'_1 Q_t)^2 + \alpha'_2 \cdot Q_t$$

If we assume that component B fails, then the conditional probability for S given B is:

$$P(S|B) = \frac{P(A \cap B)}{P(B)}$$

The system can still fail from either cut set  $\{A_i, B_i\}$  or  $\{X_{AB}\}$ . Therefore the system equation (using rate event approximation) is the sum of the following two equations:

$$P(A_i \cap B_i|B) = \frac{P(A_i \cap B_i)}{P(B)} = \frac{Q_1^{(2)} Q_1^{(2)}}{Q_T} = \frac{\alpha_1 \quad t \alpha_1 Q_t}{Q_t} = (\alpha'_1)^2 Q_t$$

$$P(X_{AB}|B) = \frac{P(X_{AB})}{P(B)} = \frac{Q_2^{(2)}}{Q_T} = \frac{\alpha_2 Q_t}{Q_t} = \alpha'_2$$

Summing these together gives:

$$\begin{aligned} P(S|B) &= P(A_i \cap B_i|B) + P(X_{AB}|B) \\ &= (\alpha'_1)^2 Q_T + \alpha'_2 \end{aligned}$$

Substituting in parameter values gives:

$$\begin{aligned}
 P(S|B) &= \alpha'_1{}^2 Q_t + \alpha'_2 \\
 &= (0.98)^2(0.006) + 0.02 \\
 &= 0.02576
 \end{aligned}$$

It is not surprising that this is the same answer as Chapter 2, given the same system equations.

For example 2, if component  $P_1$  fails, then the conditional probability for S given  $P_1$  is:

$$P(S|P_1) = \frac{P(S \cap P_1)}{P(P_1)}$$

The calculation for each cutset is shown in Table 25.

**Table 24: Cut Sets for Example 2 in event assessment**

Cut Set	$\frac{P(S \cap P_1)}{P(P_1)}$	Boolean Reduction	Basic Parameter
$\{E_{1,i}, E_{2,i}\}$	$\frac{P(E_{1,i} \cap E_{2,i} \cap P_1)}{P(P_1)}$	$P(E_{1,i} \cap E_{2,i})$	$(\alpha'_1{}^{[E]} Q_t^{[E]})^2$
$\{P_{1,i}, E_{2,i}\}$	$\frac{P(P_{1,i} \cap E_{2,i} \cap P_1)}{P(P_1)}$	$\frac{P(P_{1,i} \cap E_{2,i})}{P(P_1)}$	$\alpha'_1{}^{[P1]} \alpha'_1{}^{[E]} Q_t^{[E]}$
$\{P_{1,i}, P_{2,i}, P_{3,i}\}$	$\frac{P(P_{1,i} \cap P_{2,i} \cap P_{3,i} \cap P_1)}{P(P_1)}$	$\frac{P(P_{1,i} \cap P_{2,i} \cap P_{3,i})}{P(P_1)}$	$\alpha'_1{}^{[1]} \alpha'_1{}^{[P2]} \alpha'_1{}^{[P3]} (Q_t^{[P]})^2$
$\{E_{1,i}, P_{2,i}, P_{3,i}\}$	$\frac{P(E_{1,i} \cap P_{2,i} \cap P_{3,i} \cap P_1)}{P(P_1)}$	$P(E_{1,i} \cap P_{2,i} \cap P_{3,i})$	$\alpha'_1{}^{[E]} Q_t^{[E]} \alpha'_1{}^{[P2]} \alpha'_1{}^{[P3]} (Q_t^{[P]})^2$
$\{P_{3,i}, Z_{P1,P2}\}$	$\frac{P(P_{3,i} \cap Z_{P1,P2} \cap P_1)}{P(P_1)}$	$\frac{P(P_{3,i} \cap Z_{P1,P2})}{P(P_1)}$	$\alpha'_1{}^{[P3]} \alpha'_2{}^{[Z]} Q_t^{[P]}$
$\{E_{2,i}, Z_{P1,P2}\}$	$\frac{P(E_{2,i} \cap Z_{P1,P2} \cap P_1)}{P(P_1)}$	$\frac{P(E_{2,i} \cap Z_{P1,P2})}{P(P_1)}$	$\alpha'_1{}^{[E]} Q_t^{[E]} \alpha'_2{}^{[Z]}$
$\{P_{2,i}, Y_{P1,P3}\}$	$\frac{P(P_{2,i} \cap Y_{P1,P3} \cap P_1)}{P(P_1)}$	$\frac{P(P_{2,i} \cap Y_{P1,P3})}{P(P_1)}$	$\alpha'_1{}^{[P2]} \alpha'_2{}^{[Y]} Q_t^{[P]}$

$\{E_{2,i}, Y_{P1,P3}\}$	$\frac{P(E_{2,i} \cap Y_{P1,P3} \cap P_1)}{P(P_1)}$	$\frac{P(E_{2,i} \cap Y_{P1,P3})}{P(P_1)}$	$\alpha_1'^{[E]} Q_t^{[E]} \alpha_2'^{[Y]}$
$\{Z_{P1,P2}, Y_{P1,P3}\}$	$\frac{P(Z_{P1,P2} \cap Y_{P1,P3} \cap P_1)}{P(P_1)}$	$\frac{P(Z_{P1,P2} \cap Y_{P1,P3})}{P(P_1)}$	$\alpha_2'^{[Z]} Q_t^{[P]} \alpha_2'^{[Y]}$
$\{X_{E1,E2}\}$	$\frac{P(X_{E1,E2} \cap P_1)}{P(P_1)}$	$P(X_{E1,E2})$	$\alpha_2'^{[X]} Q_t^{[E]}$

Using rare event approximation and summing the last column of Table 25 gives  $P(S|P_1) = 6.120e-3$ . At higher levels of significant figures this estimate is slightly less than the AFM. This reduction in probability has occurred through the recognition that Pump 1 and Pump 2 are not coupled by Installation Procedure, which was not possible to account for using the AFM.

### 6.7.2. Knowledge of Failure Cause

Where the failure cause is known, the system equation can be updated using:

$$P(S|B_i) = \frac{P(A \cap B_i)}{P(B_i)}$$

where

$B_i =$  is the failure of component B due to cause  $i$ .

Given  $B_i$  has occurred, the system can now fail from the following cut sets  $\{A_I, B_{I,i}\}$  or

$\{X_{AB,i}\}$ , where the  $c$  subscript denotes the cause of the event. Therefore:

$$P(A \cap B_i) = P(A_I)P(B_{I,i}) + P(X_{AB,i})$$

The events  $B_{I,i}$  and  $X_{AB,i}$  relate to the single and CCF failure events which make up

event  $B_i = B_{I,i} \cup X_{AB,i}$ . The probability of these events can be calculated using partial alpha factors:

$$P(X_{AB,i}) = \alpha_{2,i}P(B_i)$$

$$P(B_{I,i}) = \alpha_{1,i}P(B_i)$$

Therefore:

$$\begin{aligned} P(S|B_i) &= \frac{P(A \cap B_i)}{P(B_i)} \\ &= \frac{P(A_I)P(B_{I,i}) + P(X_{AB,i})}{P(B_i)} \\ &= \frac{\alpha'_{1,i}Q_t\alpha_{1,i}P(B_i) + \alpha_{2,i}P(B_i)}{P(B_i)} \\ &= \alpha'_{1,i}\alpha_{1,i}Q_t + \alpha_{2,i} \end{aligned}$$

Table 25 shows the event assessment result for example 1 given different failure causes. This result shows that the system failure probability depends on the strength of the coupling factor to propagate failures. For example, Install Procedure has the largest partial alpha factor, which also gives the largest event assessment result.

**Table 25: Event Assessment for Example 1 with different failure causes**

<b>Cause</b>	<b><math>P(S B_C)</math></b>	<b>System Failure Probability</b>	<b><math>\alpha_{2,c}</math></b>
Unknown	$P(S B)$	0.0258	0.0200
Install Procedure Error	$P(S B_{IP})$	0.0128	0.0070
Maintenance Human Error	$P(S B_{MH})$	0.0357	0.0300
External Environment Shock	$P(S B_{EE})$	0.0624	0.0569

For example 2, if component  $P_1$  fails due to cause Maintenance Human (MH), then the conditional probability for S given  $P_{1,MH}$  is :

$$P(S|P_{1,MH}) = \frac{P(S \cap P_{1,MH})}{P(P_{1,MH})}$$

The calculation for each cut set is shown in Table 26.

**Table 26: Cut Sets for Example 2 in event assessment**

Cut Set	$\frac{P(S \cap P_{1,MH})}{P(P_{1,MH})}$	Boolean Reduction	Basic Parameter
$\{E_{1,i}, E_{2,i}\}$	$\frac{P(E_{1,i} \cap E_{2,i} \cap P_{1,MH})}{P(P_{1,MH})}$	$P(E_{1,i} \cap E_{2,i})$	$(\alpha_1^{[E]} Q_t^{[E]})^2$
$\{P_{1,i}, E_{2,i}\}$	$\frac{P(P_{1,i} \cap E_{2,i} \cap P_{1,MH})}{P(P_{1,MH})}$	$\frac{P(P_{1,i,MH} \cap E_{2,i})}{P(P_{1,MH})}$	$\alpha_{1,MH}^{[P1]} \alpha_1^{[E]} Q_t^{[E]}$
$\{P_{1,i}, P_{2,i}, P_{3,i}\}$	$\frac{P(P_{1,i} \cap P_{2,i} \cap P_{3,i} \cap P_{1,MH})}{P(P_{1,MH})}$	$\frac{P(P_{1,i,MH} \cap P_{2,i} \cap P_{3,i})}{P(P_{1,MH})}$	$\alpha_{1,MH}^{[P1]} \alpha_1^{[P2]} \alpha_1^{[P3]} (Q_t^{[P]})^2$
$\{E_{1,i}, P_{2,i}, P_{3,i}\}$	$\frac{P(E_{1,i} \cap P_{2,i} \cap P_{3,i} \cap P_{1,MH})}{P(P_{1,MH})}$	$P(E_{1,i} \cap P_{2,i} \cap P_{3,i})$	$\alpha_1^{[E]} Q_t^{[E]} \alpha_1^{[P2]} \alpha_1^{[P3]} (Q_t^{[P]})^2$
$\{P_{3,i}, Z_{P1,P2}\}$	$\frac{P(P_{3,i} \cap Z_{P1,P2} \cap P_{1,MH})}{P(P_{1,MH})}$	$\frac{P(P_{3,i} \cap Z_{P1,P2,MH})}{P(P_{1,MH})}$	$\alpha_1^{[P3]} Q_t^{[P]} \alpha_{2,MH}^{[Z]}$
$\{E_{2,i}, Z_{P1,P2}\}$	$\frac{P(E_{2,i} \cap Z_{P1,P2} \cap P_{1,MH})}{P(P_{1,MH})}$	$\frac{P(E_{2,i} \cap Z_{P1,P2,MH})}{P(P_{1,MH})}$	$\alpha_1^{[E]} Q_t^{[E]} \alpha_{2,MH}^{[Z]}$
$\{P_{2,i}, Y_{P1,P3}\}$	$\frac{P(P_{2,i} \cap Y_{P1,P3} \cap P_{1,MH})}{P(P_{1,MH})}$	$P(Y_{P1,P3} \cap P_{1,MH}) = 0$	0
$\{E_{2,i}, Y_{P1,P3}\}$	$\frac{P(E_{2,i} \cap Y_{P1,P3} \cap P_{1,MH})}{P(P_{1,MH})}$	$P(Y_{P1,P3} \cap P_{1,MH}) = 0$	0
$\{Z_{P1,P2}, Y_{P1,P3}\}$	$\frac{P(Z_{P1,P2} \cap Y_{P1,P3} \cap P_{1,MH})}{P(P_{1,MH})}$	$\frac{P(Z_{P1,P2,MH} \cap Y_{P1,P3})}{P(P_{1,MH})}$	$\alpha_{2,MH}^{[Z]} \alpha_2^{[Y]} Q_t^{[P]}$
$\{X_{E1,E2}\}$	$\frac{P(X_{E1,E2} \cap P_{1,MH})}{P(P_{1,MH})}$	$P(X_{E1,E2})$	$\alpha_2^{[X]} Q_t^{[E]}$

Using rare event approximation and summing the last column of Table 26 gives  $P(S|P_{1,MH}) = 6.053e-3$ , which is slightly lower than the estimate without causes,  $P(S|P_1) = 6.120e-3$ , because MH has a lower propagation probability,  $\alpha_{2,MH}$  than the general alpha factor  $\alpha_2^{[Y]}$ . Table 27 shows the probability of system failure for each cause of Pump 1.



**Table 27: Event Assessment for Example 2 with different failure causes**

<b>Cause</b>	$P(S P_{1,C})$	<b>System Failure Probability</b>
Unknown	$P(S P_1)$	6.120e-3
Install Procedure Error	$P(S P_{1,IP})$	6.100e-3
Maintenance Human Error	$P(S P_{1,MH})$	6.053e-3
External Environment Shock	$P(S P_{1,EE})$	6.154e-3

## 6.8. Data Collection Requirements

### 6.8.1. *Desirable Data Collection Attributes*

In order to conduct impact vector analysis with the partial alpha factor model, the following data collection attributes are desirable. Where this data is not currently available within the CCFDB, the assumption required to use the CCFDB is stated.

#### **The failure cause is recorded for single and multiple failure events**

In order for PAFM model parameter estimates to be based on data, the failure cause for single and multiple failure events is required. This is currently available in the CCF database.

**The potential coupling factors through which failure propagation could occur is recorded.** Data which makes up the failure event databases are recorded from real systems which have varying degrees of coupling factors between them. Therefore it would be ideal if a failure database system recorded the potential coupling factors through which a failure could have propagated, for each recorded failure. This would

require a significant increase in the resources required to record failure data.

Where this data is unavailable, the following assumption is made:

***PAFM Data Assumption 1:** All components within the CCCG are susceptible to all failure causes and have perfect symmetry within the CCCG. That is to say that failure causes have an equal probability of propagation to each component within the CCCG. This assumption is already used in the Basic Parameter Model and the Alpha Factor Model.*

**A mutually exclusive, one to one relationship between failure causes and coupling factors**

The previous data collection attribute is unlikely to be economical to achieve. Therefore a failure data taxonomy was proposed in Chapter 4, to minimize the impact of assumptions required to overcome this data deficiency.

During the analysis of historic data, single failures have very little information about the possible propagation paths for the failure, other than the failure cause. Where a one to one relationship between failure cause and coupling factor exists, an assumption may be made that the single failure had the potential to propagate only through its paired coupling factor. Furthermore, where a cause can propagate through multiple coupling factors, the discriminatory ability of the partial alpha factor model to quantify event assessments based on coupling factors is reduced.

Using this data taxonomy introduces the following assumption:

***Assumption 2.** Given knowledge of a failure cause, the coupling factors over which the cause could propagate are known.*

The occasion when a one to one relationship does not exist will be discussed in section 6.8.2.

### **Size of Common Cause Component Groups for Single Failures**

Currently the size of the CCCG for single failures is not recorded. This information is required for the construction of impact vectors during the parameter estimation step. Where the CCCG size for the single failure is different to the target system, the impact vector is required to be mapped to the new size, as detailed in section 0. Table 28 shows how the impact vectors for target system can change dramatically when using different assumptions of CCCG size for an observed single failure.

**Table 28: Demonstration of impact vectors changing for target system based on assumption of CCCG size for observed single failure**

<b>Assumed CCCG size for observed failure event</b>	<b>Impact Vector of Single Failure</b>	<b>Equivalent impact vector after being mapped to CCCG size of 4</b>
1	[1]	[4, 0, 0, 0]
2	[1,0]	[2, 0, 0, 0]
3	[1,0,0]	[1.33, 0, 0, 0]
4	[1,0,0,0]	[1, 0, 0, 0]
5	[1,0,0,0,0]	[0.8, 0, 0, 0]
6	[1,0,0,0,0,0]	[0.66, 0, 0, 0]
7	[1,0,0,0,0,0,0]	[0.57, 0, 0, 0]

Currently this quantity is estimated as an average of CCCG sizes recorded on the multiple failure CCF events (Wierman & Kvarfordt 2011).

Where this data is unavailable, the following assumption is made:

***PAFM Data Assumption 3:** The size of the CCCG for each single failure event is the same as the average CCCG size for multiple failure events.*

**6.8.2. When a one to one relationship between cause and coupling factor does not exist.**

The following section will demonstrate the use of the PAFM where a failure data taxonomy allows for a failure cause to propagate through multiple coupling factors.

Table 29 shows the failure data for a system. The items in red text are data which is not currently found within the CCFDB. Of particular note is the addition of a column “Potential CF” which is an assessment of which coupling factors that particular failure could propagate through.

**Table 29: Example data when a failure cause can propagate through multiple coupling factors**

#	No. Fail	Pop Size	Cause	Potential CF	Coupling Factor	Impact Vector	
						F1	F2
1	1	2	PC 1	CF1		1	0
2	1	2	PC 2	CF1, CF2		1	0
3	2	2	PC 3	CF2	CF2	0	1
4	1	2	PC 1	CF1, CF2		1	0
5	1	2	PC 3	CF1		1	0
6	2	2	PC 1	CF1, CF2	CF2	0	1
7	1	2	PC 2	CF1, CF2		1	0
8	1	2	PC 2	CF1		1	0
9	2	2	PC 1	CF1	CF1	0	1
10	2	2	PC 1	CF1, CF2	CF2	0	1
TOTAL						6	4

This example shows that of all the observed failure events, four were CCF of two components, and six were single failures. Therefore:

$$\alpha_1 = 0.6, \quad \alpha_2 = 0.4$$

The partial alpha factors can be estimated as:

$$\alpha_{2,CF1} = \frac{n_{2,CF1}}{n_{p,CF1}} = \frac{1}{9}$$

$$\alpha_{2,CF2} = \frac{n_{2,CF2}}{n_{p,CF2}} = \frac{3}{6}$$

The gamma factors can be estimated as:

$$\gamma_{CF1} = \frac{n_{p,CF1}}{n_t} = \frac{9}{10}$$

$$\gamma_{CF2} = \frac{n_{p,CF2}}{n_t} = \frac{6}{10}$$

The assessed alpha factor for a symmetrical two train system is:

$$\begin{aligned}\alpha_2 &= \gamma_{CF1}\alpha_{2,CF1} + \gamma_2\alpha_{2,CF2} \\ &= \frac{9}{10} \cdot \frac{1}{9} + \frac{6}{10} \cdot \frac{1}{2} \\ &= 0.4\end{aligned}$$

This simple example shows an application of the PAFM where a cause has the potential to propagate through multiple coupling factors. The estimates for the partial alpha factor and gamma factor were able to be obtained, if given the information about the potential coupling factor avenues for each observed failure. When used in this way, the gamma factors do not sum to 1, and are not distributed with a multinomial distribution. The ability for the PAFM parameters to estimate the system alpha factors was demonstrated.

### ***6.9. Model Assessment***

The Partial Alpha Factor Model provides one additional level of detail to include in the assessment of common cause failures compared to the AFM. It does however use the same empirical ratio methods of the alpha factor model and therefore most of the issues which exist with the current methodology remain. The PAFM parameters can be used to assist in the quantification of the General Dependency Model parameters.

### ***6.9.1. Model Advantages***

The aim of the PAFM is to extend the AFM such that event assessments can be conducted with knowledge of the failure cause. Given quantified parameters, the PAFM model achieves this with minimal changes to the AFM methodology.

Specifically the advantage of the PAFM over other CCF models are summarized as:

- Allows greater resolution on event assessments.
- Intuitive extension to the AFM analysis methodology.
- A ratio model allowing the use of target system failure rates.
- Can reward target system defenses that decouple dependencies.
- Can use AFM for system analysis, and the PAFM for event assessment.
- The PAFM can be calculated from the CCFDB.
- PAFM parameter estimates will be no worse than if the PRA used the AFM.

### ***6.9.2. PAFM Limitations***

Due to the PAFM using an AFM methodology and due to the nature of CCF data the PAFM has a number of limitations:

- The description of the target system features such as cause and coupling factor features and defenses is limited.

- Many of the failure causes will have no observed CCF events and therefore the parameter estimates rely more on the prior knowledge.
- As per the AFM, it is difficult to model components with different failure probabilities within the same CCCG (symmetrical failure probabilities)
- In order to use the CCFDB, it must be assumed that each component within the CCCG for the observed failure has the potential for propagation of that cause through a coupling factor. This assumption may not be true and will produce an optimistic estimate.
- Impact vector mapping is still required if data is from a different size CCCG.

### ***6.9.3. Compare Against Model Criteria***

Table 30 provides a comparison of the PAFM features compared to previously proposed models.



**Table 30: Assessment of the PAFM compared to previous CCF models**

	Partial Alpha Factor Model	Basic Parameter	Beta Factor	Partial Beta Factor	Alpha Factor Model	Binomial Failure Rate Model with Lethal Shocks	Common Load	Reliability Cut Off	Influence Diagram	Bayesian Network
Feature Description	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Explicitly Models System Features	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Models failure cause	Y	N	N	P	N	N	N	P	Y	Y
Models failure cause defense	N	N	N	Y	N	N	N	Y	N	P
Models coupling factor	Y	N	N	P	N	N	N	P	Y	N
Models coupling factor defense	P	N	N	Y	N	N	N	Y	N	N
Models deeper causal levels	N	N	N	N	N	N	N	N	N	Y
Models cause condition / shock	N	N	N	N	N	Y	Y	N	N	Y
Models multiplicity of failures within CCCG	Y	Y	N	N	Y	Y	Y	N	N	Y
Models includes consideration for rectification period	N	N	N	N	N	N	N	N	N	N
Common Cause Component Grouping Characteristics	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Model non-symmetrical but similar components within the same CCCG	Y	N	N	N	N	N	N	N	N	Y
Model different components within the same CCCG	N	N	N	N	N	N	N	N	N	Y
A component can be part of many CCCGs	Y	N	N	N	N	N	N	N	N	Y
No limit to CCCG size	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Model different failure multiplicities within the CCCG ( $k$ failures in $n$ )	Y	Y	N	N	Y	Y	Y	N	N	Y

Event Assessment Capabilities	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Event Assessment with knowledge of a failed component	Y	Y	N	N	Y	Y	?	N	Y	Y
Event Assessment with knowledge of failure cause	Y	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence - Partial Failures	N	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence- Virtual evidence of cause	N	N	N	N	N	N	N	N	Y	Y
Parameter Estimation	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Impact Vector Method (including method for incorporating uncertainty)	Y	Y	P	N	Y	Y	N	N	N	N
Expert estimations (in absence of any data)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account for reliability growth (discount previous failures)	N	N	N	N	N	N	N	N	N	N
Update parameters with new evidence	Y	Y	P	N	Y	Y	Y	N	N	N
Incorporate evidence from different sized CCCGs	Y	N	P	N	P	Y	Y	N	N	N
Account for CCF which occurred in a different mission time	N	N	N	N	N	N	N	N	N	N
Account for CCF data which has artificial separation in time due to	N	N	N	N	N	N	N	N	N	N
Use system specific failure rate data combined with generic model	Y	N	Y	N	Y	N	N	N	N	N
Uncertainty Characteristics for Parameter Estimation	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Does not require distinguish between independent and single CCF failures	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Failures outside the mission period	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty of shared cause	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty of coupling factor	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty in intervals due to staggered testing	P	P	P	N	P	P	N	N	N	N
Partial failures and component degradation	Y	Y	P	N	Y	Y	N	N	N	N
Usability and Cultural Considerations	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Backward compatible to Alpha Factor Model parameters	Y	Y	N	N	Y	N	N	N	N	N
The time investment is no more than the alpha factor model.	Y	Y	Y	Y	Y	Y	N	Y	N	N
Automatic parameter estimation is possible from the CCFDB/RADs	Y	Y	Y	N	Y	Y	N	N	N	N

## Chapter 7: General Dependency Model

### 7.1. Introduction

#### 7.1.1. *Motivation*

The motivation for the General Dependency Model is to achieve the research objectives without consideration for the constraints of the current methodology.

The objectives for the General Dependency Model are to:

- Enable event assessment with knowledge of the failure event's characteristics.
- Model the increased and decreased propensity of a system to experience CCF based on the system features such as causes, coupling factors and defenses.
- Model asymmetrical dependency relationships.
- Model asymmetrical components.
- Retain the modeling of different multiplicities of failures.
- Allow for parameter estimation using the impact vector methodology.
- Allow an analysis procedure which is no more complex than using the AFM.

The General Dependency Model design has been motivated by features of existing CCF models covered in Chapter 5. A brief summary of those motivations are provided below.

**Influence Diagram Model and Event Assessment Model** (Kelly et al. 2011). Both of these models are based on a Bayesian Network, which has also been chosen for the General Dependency Model for the following reasons:

- The Bayesian Network can show complex causal relationships through an intuitive graphical representation.
- The Bayesian Network can model soft dependencies between random variables (as opposed to deterministic dependencies which fault trees and event trees are limited to).
- The Bayesian Network allows a complex joint probability distribution to be quantified through local probability relationships.
- The Bayesian Network model is an advanced field of study which already has mathematical algorithms and software packages to calculate the relationship between random variables and evidence propagation. This allows this thesis to focus on model construction without the need for complex details to demonstrate the model's capabilities.

**Unified Partial Method** (Brand & Gabbot 1993). UPM is a combination of the Partial Beta Factor Model (Johnston 1987) and the Reliability Cut of Method (Bourne et al. 1981). The essential feature of both methods is a subjective assessment of the defenses and dependencies which exist within the system. An essential feature of the GDM is to encode the specific results found within the qualitative assessment of the system into the model structure (see section 2.4.1)

**Shock Models.** Shock models explicitly model the cause condition (shock) to the system and model a failure through a fragility parameter,  $p_i$ , in the presence of a shock. This allows the modeling of any level of redundancy without having data from systems with the same CCCG sizes. The GDM model also models failure based on the cause condition frequency and component fragility.

The shock model approach has been criticized for inaccuracy at high levels of redundancy. This is because of the assumption that all components receive the same shock. The GDM model assumes that the propagation of the shocks to each component is probabilistic and this assumption can be controlled through use of the GDM parameters. Furthermore the shock models assume perfect symmetry of components which may be unlikely for large CCCGs. GDM accounts for asymmetrical features of large CCCGs.

**Ratio models.** Ratio models such as the Beta Factor (Fleming 1975) provide an empirical relationship between common cause failure and single failures. GDM uses a similar concept, except instead of the empirical ratio being used to model failure propagation, it is used to model cause condition propagation. For example, the probability that a cause condition is present on multiple components is modeled using a ratio metric. By modeling the cause condition instead of the failures, the GDM approach avoids the need to distinguish between a single failure and an independent

failure.

The PAFM (chapter 6) uses ratio parameters for each failure cause. GDM will use the PAFM values as an intermediary step to estimate its parameters. This provides the ability to quantify GDM parameters using industry data and ensure the GDM is consistent with the AFM estimates.

### **7.1.2. Chapter Scope**

This chapter will discuss:

- An overview of the model
- An overview of Bayesian Networks
- Description of the model structure
- Description of the model parameters.
- Formulate the parameter estimation equations
- Describe methods to quantifying parameters.
- Describe a system analysis method.
- Describe an event assessment method.
- Describe the data requirements for the model.
- Summarize and evaluate the model.
- Describe extension and future work required to implement the model.

### 7.1.3. Examples

In order to demonstrate CCF analysis using the General Dependency Model, the same two examples used in Chapter 2 and Chapter 6 will be used. Example 1 is a two train example with identical EDGs. Example 2 is a mixed redundancy system consisting of two EDGs and three pumps.

## 7.2. Model Structure

### 7.2.1. Component Failure Probability

The General Dependency Model defines the component failure rate,  $Q_t$  is the combination of component failure probabilities for each failure cause. To the component, each cause is independent of each other and using rare event approximation the component failure probability can be calculated as:

$$P(A) = Q_t = \sum_{i=1}^w Q_{t,i}$$

- $A$  = A random variable for the failure of component A  
 $Q_t$  = The total failure probability for a component  
 $Q_{t,i}$  = The failure probability of a component due to cause  $i$ .

The ‘cause’ is defined as a cause condition from which a failure can occur. The probability that a cause condition exists for cause  $i$  is:

$$P(C_i) = Q_{E,i}$$

- $C_i$  = A random variable for the existence of cause condition  $i$ .  
 $Q_{E,i}$  = The cause condition probability of cause  $i$ .

In the presence of an cause condition,  $C_i$ , the probability that the component fails is  $p_i$ .

Therefore the probability of component failure due to cause  $i$  is:

$$Q_{t,i} = p_i Q_{E,i}$$

$p_i =$  *the probability a component fails when tested by cause  $i$ .*

The failure of component  $A$ , is the union of contributions from each failure cause. Using rare event approximation this equals:

$$P(A) = \sum_{i=1}^w p_i Q_{E,i} \text{ (rare event approximation)}$$

Without the assumption of rare event approximation,  $P(A)$  can be calculated using:

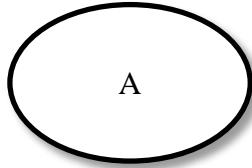
$$\begin{aligned} P(A) &= 1 - \prod_{i=1}^w (1 - Q_{t,i}) \\ &= 1 - \prod_{i=1}^w (1 - p_i Q_{E,i}) \end{aligned}$$

This concept is shown in Figure 31.



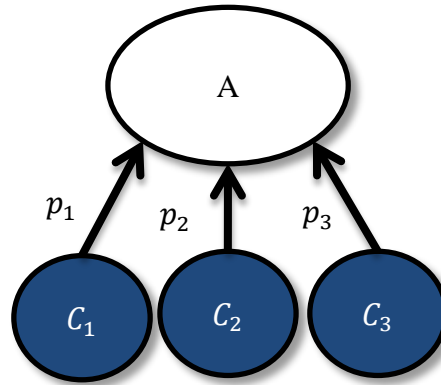
### Traditional Basic Event

$$P(A) = Q_T$$



### GDM Basic Event

$$P(A) = 1 - \prod_{i=1}^w (1 - p_i Q_{E,i})$$



$$P(C_2) = Q_{E,2}$$

$$P(C_1) = Q_{E,1}$$

$$P(C_3) = Q_{E,3}$$

Figure 31: GDM Basic Events

#### 7.2.2. Component Dependency

Thus far the model has included the cause conditions and failure probabilities which are local to a component. It is possible that multiple components share the same cause condition due to a coupling factor. The presence of coupling factors between components is identified during the qualitative assessment of the target system features (see section 2.4.1). By coupling the cause condition, instead of failure causes, the model can now better describe the physical phenomena of CCF. It may be possible for a cause condition to affect many different components (like extreme temperatures). By separating the cause condition, and the response of the component to that cause condition, it is easier to model asymmetrical relationships.

Figure 32 shows an EDG and pump which share the same location. If the EDG suffers from an extreme environmental condition, then the pump will also experience the same condition (or shock). The difference between the EDG and pump in the presence of such a cause condition, is the fragility of each component to withstand the shock,  $p_i$ .

### GDM Basic Event

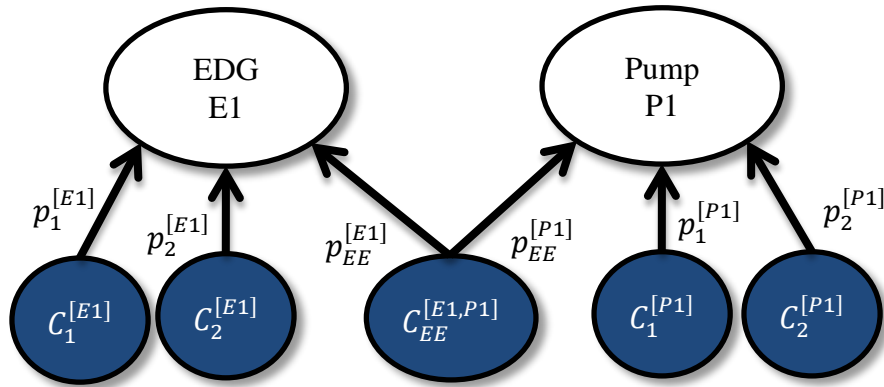


Figure 32: GDM Coupling Components

#### 7.2.3. Propagation of Cause Condition

The model, thus far, has assumed the propagation of a cause condition to other components is certain, where a coupling factor exists. However the propagation of a cause is likely to be probabilistic. For example, a maintenance tradesman is inexperienced and conducts a maintenance error on an EDG. The tradesman progresses to maintain a second EDG. Despite the two components being coupled by the same maintainer, the likelihood of the second EDG suffering the same maintenance error is probabilistic.

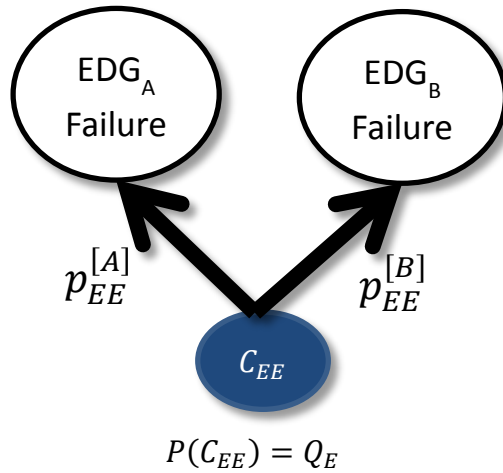
This model structure is problematic because the model:

- cannot model scenarios where components have high fragility to an cause condition, and a low probability of CCF, and
- cannot account for defenses against cause condition propagation such as protecting against external environmental cause conditions by moving components into separate rooms.

The GDM separates the local cause condition for each component. Local cause conditions can propagate to other components probabilistically using a coupling strength factor,  $\eta_i$ . Figure 33 shows conceptually how the GDM model can account for high fragility and low coupling factor strength through the use of local cause condition nodes.

### High Fragility

1. EDG<sub>A</sub> fails
2. High probability of EE
3. High probability of EDG<sub>B</sub> failure
4. High  $\alpha_2$



### High Fragility With CF Defense

1. EDG<sub>A</sub> fails
2. High probability of EE<sub>A</sub>
3. Low probability of EE<sub>B</sub>
4. Low probability of EDG<sub>B</sub> failure
5. Low  $\alpha_2$

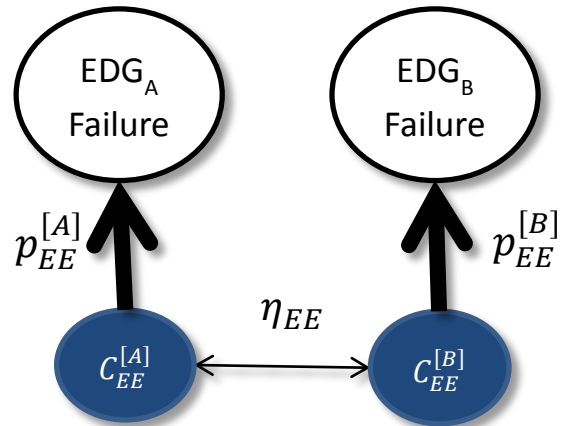


Figure 33: Conceptual propagation of cause condition through coupling factor

The coupling factor strength,  $\eta_i$ , needs to scale between the following two extremes:

- $\eta_i = 0$ . When the coupling factor strength is zero, there is no chance that the local cause condition at one component can propagate to the second component. However the second component may still fail from an independent occurrence of that cause condition.
- $\eta_i = 1$ . An cause condition at either component means that the same cause condition is present at the other component.

In order to model these limits, the cause condition probability,  $Q_{E,i}$  is split into

independent ( $Q_{IE,i}$ ) and common error ( $Q_{CE,i}$ ) probabilities.

$$P(X_i) = Q_{CE,i} = \eta_i Q_{E,i}$$

$$P(I_i) = Q_{IE,i} = (1 - \eta_i) Q_{E,i}$$

$X_i =$  *random variable for the common cause condition for cause i.*

$I_i =$  *random variable for the independent cause condition for cause i.*

$\eta_i =$  *the coupling factor strength for cause i.*

The common cause condition and independent cause condition are mutually exclusive events. Therefore the local cause condition probability is the sum of the independent and common cause condition probabilities.

$$C_i = I_i \cup X_i$$

$$Q_{E,EE} = Q_{IE,EE} + Q_{CE,EE}$$

$C_i =$  *A random variable for the existence of cause condition i.*

$Q_{E,i} =$  *The cause condition probability of cause i.*

Figure 33 shows the construction of the GDM model with consideration for a coupling factor strength parameter.

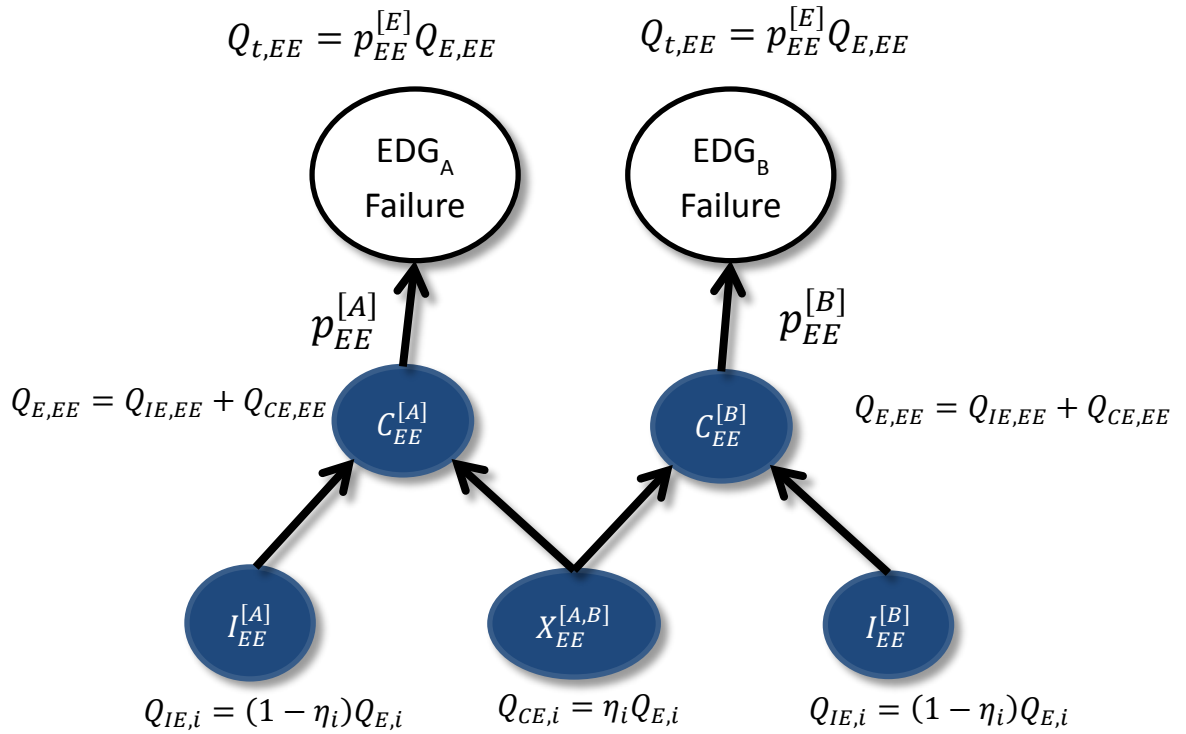


Figure 34: Conceptual construction of the GDM model

#### 7.2.4. Parameter Description

A strength of GDM is the relationship between the model parameters and the features of the target system. This subsection will review how the GDM parameters should be interpreted.

For each failure cause classification, the GDM has three parameters, the component fragility, the cause condition probability, and the coupling factor strength.

**Fragility.**  $p_i$  is the probability a component will fail given that a cause condition is

evident for cause  $i$ . It is a measure of the components ability to resist failure. The component's fragility is affected by such things as the component's design, materials, derating and compliance to reliability durability standards.

**Cause Condition Probability.**  $Q_{E,i}$  is the probability that the Cause condition for cause  $i$  is present. The cause condition probability represents the frequency and strength of failure causes. It is a function of features such as quality assurance, process maturity, human performance shaping factors. This parameter is similar to the Binomial Failure Rate Model's, rate of shocks. The term has been renamed to recognize that a cause condition may exist for extended periods of time, and therefore may not be considered a shock. This consideration will be discussed in more detail below.

**Coupling Factor Strength.**  $\eta_i$  is the probability that if an cause condition exists at a component, that it will be propagated to other components. The coupling factor strength is a measure of defenses against coupling factors, and of the repeatable nature of the cause condition.

Central questions for the physical interpretation of the GDM parameters are:

- Can multiple cause conditions (shocks) test the system during a single mission period?
- Is the term  $Q_{E,i}$  a rate or a probability?
- Can a cause condition last over multiple mission periods?
- Can multiple cause conditions exist in the one cause category at the same time, and does this increase the probability of failure?

These are important questions, which can be analyzed in context of the interpretations taken by previously proposed CCF models. First the outcome of the model must be restated.

### **Multiple Events Occurring During A Mission Period**

A reasonable question, when relating the CCF model to the physics of the phenomena is whether multiple cause conditions (or shocks), of the same cause, may exist during a single mission period. Examples of this may be multiple maintenance procedural errors occur during a service, or when multiple independent installation errors occurred on a component. How does the model interpret such events?

PRA models quantify the probability of failure during a mission period or average unavailability. In order to do this, each basic event must quantify the probability of failure for the mission period or the average unavailability. All CCF models to date (see appendix 1), have proposed CCF as a mission reliability figure, without consideration for repair (see section 2.2 and 3.4.2). The investigation of GDM to consider repair times, and hence provide an unavailability probability is left to future research (see section 8.7.1).

The term  $Q_{t,i}$  quantifies the probability of a component to fail during the mission. The mission is modeled as a single event, and therefore the model does not allow for a



component cannot fail multiple times during a single mission. From an interpretation point of view, the analyst can only observe the result at the end of the mission period, where the question is asked, “Did the component fail due to cause  $i$ ”. The answer to this question can only be a single failure or no failure.

The same interpretation must be used when using impact vectors. As discussed in section 3.4.2, the definition of a common cause failure, for the purposes of creating impact vectors must include the mission time. Therefore the impact vector is equivalent to an observation at the end of a mission period where the question is asked, “How many components in the CCCG failed due to a shared cause?”. The answer provides no insight if multiple shocks occurred during the mission, only the outcome at the end.

The cause condition term,  $Q_{E,i}$ , is also required to conform to the interpretation restrictions of  $Q_{t,i}$  and the impact vector. The mission period is considered a single event, after which we ask, “Was a cause condition present?”. Therefore the PRA model cannot explicitly account for multiple shocks occurring. On occasions where multiple shocks are common, this will be modeled through either an increase in the cause condition rate,  $Q_{E,i}$  or through an artificial decrease in the fragility term,  $p_i$ .

A mathematical solution to overcome this problem will be discussed at the end of the section ‘Cause Condition As a Rate’. This solution will not be implemented as part of this thesis due to a reliance on the assumptions used to create impact vectors.

## Cause Condition As a Rate

In conducting probability modeling, it is important to distinguish between a rate of occurrence, and a probability. A rate has a dimension such as, ‘per demand’, or ‘per hour’, while a probability is dimensionless and ranges between 0 and 1.

During Chapter 2 the basic parameter term  $Q_k^{(m)}$  was referred to as a probability or rate.

(Mosleh et al. 1998) defines  $Q_k^{(m)}$  as a probability, while defining  $Q_t$  as a failure frequency (a rate). The relationship between the two terms is:

$$Q_k^{(m)} = \binom{m-1}{k-1}^{-1} \cdot \alpha_k \cdot Q_t \quad \text{staggered test data}$$

The first term is a combination calculation and dimensionless. The  $\alpha_{ki}$  term is a ratio metric and also dimensionless. So if  $Q_t$  has a dimension, then so must  $Q_k^{(m)}$ . Further confusion can be provided through the definition of the basic parameter term and the Binomial Failure Rate Model (Atwood 1986):

$$Q_k^m = \begin{cases} Q_I + \mu \cdot \rho(1 - \rho)^{m-1} & \text{where } k = 1 \\ \mu \cdot \rho^k(1 - \rho)^{m-k} & \text{where } 2 \leq k < m \\ \mu \cdot \rho^m + \omega & \text{where } k = m \end{cases}$$

The terms  $\mu$  and  $Q_I$  are the rate of common shocks, and the independent failure rates respectively. The remaining terms involving  $\rho$  are the binomial distribution resulting in a probability which is dimensionless. In essence, the formula multiplies a shock rate with a probability of failure to get a failure rate, which is similar to the GDM

calculation,  $Q_{t,i} = p_i Q_{E,i}$ .

What is the effect of defining failure probabilities using rates?

To remove ambiguity from the following discussion lets define  $\lambda_{E,i}$  as the cause condition (shock) rate, and  $Q_{E,i}$  as the probability that an error existed during the mission period,  $t_M$ . Assuming the rate of each occurrence is exponentially distributed between arrival times, at a constant rate,  $\lambda_{E,i}$ , then the probability of receiving  $k$  shocks during a mission period is calculated using the Poisson distribution:

$$P(K = k) = \frac{(\lambda_{E,i} t_M)^k}{k!} e^{-\lambda_{E,i} t_M}$$

Where the rate parameter is in the same dimensions as the mission period (i.e a 'per 24 hours' rate where  $t_M = 24hrs$ ) the Poisson distribution becomes:

$$P(K = k) = \frac{\lambda^k}{k!} e^{-\lambda_{E,i}}$$

As discussed in the previous section, the term  $Q_{E,i}$  is the probability that one or more shocks occurred during the mission period. Therefore the cause condition rate can be converted to a mission probability value using:

$$\begin{aligned}
 Q_{E,i} &= P(K > 1) = 1 - P(K = 0) \\
 &= 1 - e^{-\lambda_{E,i}}
 \end{aligned}$$

Table 31 shows a comparison of  $\lambda_{E,i}$  and  $Q_{E,i}$  terms to see where they can be used interchangeably. It shows that for rare events (less than 0.001) the terms approximately equal each other,  $\lambda_{E,i} \approx Q_{E,i}$ .

**Table 31: Comparison of probability and rate metrics**

$\lambda_{E,i}$	$Q_{E,i}$
100	1
10	0.999955
1	0.632121
0.1	0.095163
0.01	0.00995
1E-03	1.00E-03
1E-04	1.00E-04
1E-05	1.00E-05
1E-06	1.00E-06

Therefore the interchangeability of rates and probabilities which has occurred within the CCF literature uses a rare event approximation assumption. For the purposes of this thesis,  $Q_{E,i}$  will represent the probability a cause condition occurs during the mission period, and not a rate metric. This is because the rare event approximation assumption

may not be true, as will be seen in section 7.6.

With the Poisson distribution being introduced, the answer to the question of  $Q_{t,i}$  modeling multiple shocks may now be answered. The Poisson distribution provides a probability value for each possible number of shocks which may occur during a mission period. If two shock occur, then the probability of failure is  $p_i + p_i - p_i^2$  or  $1 - (1 - p_i)^2$ . So a general formula for modeling  $Q_{t,i}$  with multiple shocks is:

$$Q_{t,i} = \sum_{j=1}^{\infty} \{1 - (1 - p_i)^j\} \left\{ \frac{(\lambda_{E,i})^j}{j!} e^{-\lambda_{E,i}} \right\}$$

This additional complexity in modeling multiple shocks will not be considered further in this thesis, as the quantification methods use the impact vector methodology which assumes either a cause condition existed or not.

### **Cause Conditions with Extended Durations**

Another reasonable question regarding physical interpretation of the GDM parameter  $Q_{E,i}$  regards the concept that an cause condition may have existed for multiple mission periods. An example a installation error may have been in existence for years before a failure was observed. How does the GDM interpret such an event?

The parameter  $Q_{E,i}$  has been defined as the probability an cause condition exists during a mission period, regardless of how many shocks occur during the mission, nor with regard to whether the condition originated during that mission. In this regard,  $Q_{E,i}$  may represent a ratio of mission periods with a cause condition over the total mission periods. Using this interpretation it is easy for  $Q_{E,i}$  to approach 1, especially for long duration cause conditions such as installation procedural errors. Therefore the assumption in the rare event approximation used in the previous section to use  $\lambda_{E,i} \approx Q_{E,i}$  may not be true.

### **Parameter Interpretation and Comparison**

The preceding sections have discussed the interpretation of the GDM parameters in relation to the mechanics of how CCFs occur. The model parameters represent discrete parts of the CCF phenomena, modeling cause strength, coupling factor strength and component fragility. However due to limitations in modeling techniques use in PRA, impact vectors and the GDM model itself the parameters may become more abstract, as per the other CCF models. Therefore the model parameters may only form an indication of its intended representation.

### **7.3. Bayesian Networks**

This section will provide an overview of Bayesian Networks to enable an understanding of the construction of the General Dependency Model. The section will

attempt to be brief, and focus the features of a Bayesian Network.

### **Bayesian Network Example**

To expedite the introduction of Bayesian Networks, a famous example of a house burglary will be used (Pearl 1988). The problem statement is (Korb & Nicholson 2004):

*“You have a new burglar alarm installed. It reliably detects burglary, but also responds to minor earthquakes. Two neighbors, John and Mary, promise to call the police when they hear the alarm. John always calls when he hears the alarm, but sometimes confuses the alarm with the phone ringing and calls then also. On the other hand, Mary likes loud music and sometimes doesn’t hear the alarm. If an earthquake occurs it is likely to be reported on the radio news. Given evidence about who has and hasn’t called and a radio report, you’d like to estimate the probability of a burglary”*

The Bayesian Network for this example is shown in Figure 35.

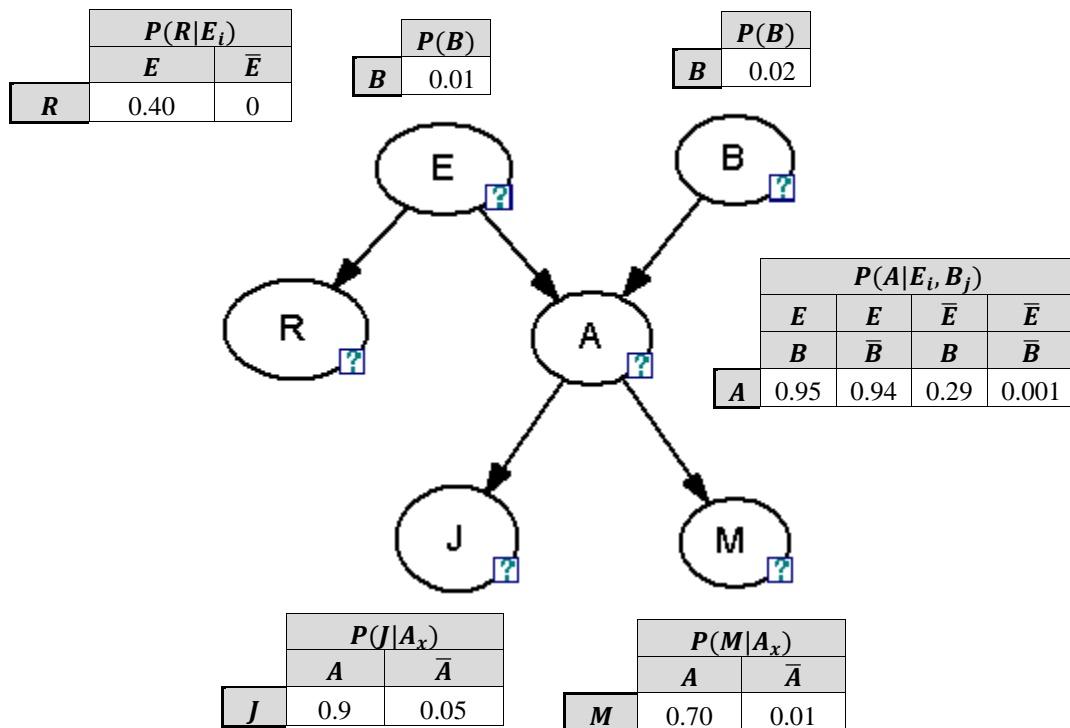


Figure 35: Bayesian Network for Burglary Example

### 7.3.1. Bayesian Network Components

Bayesian Networks are a graphical structure for representing the probabilistic relationships among a large number of variables and doing probabilistic inference with those variables (Neapolitan 2003).

A Bayesian Network consists of a number of components:

- Nodes and values (Random variables)
- Structure (Dependencies)



- Conditional Probabilities (Conditional Probability Tables)

*Nodes* represent random variables of event outcomes. Events may represent predicted events, explanatory events or evidence variables. For the burglary example, the random variables represent the following events:

E: Earthquake occurs  
B: Burglary occurs  
A: Alarm occurs  
J: John calls  
M: Mary calls  
R: Radio reports earthquake

Each random variable has distribution over the possible outcomes for the event called states. The states for a node are required to be mutually exclusive and exhaustive. For the burglary example, each node has two states {True, False}.

*Structure* captures the qualitative dependency between variables. These are represented as arcs and can be interpreted as having a causal direction, or simply a node being conditional on another node. For the burglary example an Earthquake or Burglary may cause the alarm to sound. Therefore links from the earthquake node (E) and burglary node (B) are directed to the alarm node (A).

*Conditional Probability* quantifies the dependency relationship between the nodes. This is done by populating the Conditional Probability Table for each node.

For example, the events, Earthquake and Burglary are parent nodes with no conditional events. They have a probability of occurring of 0.01 and 0.02 respectively. The probability of John calling is dependent on whether an alarm actually occurred. If the alarm sounded, then its highly likely that John would have called ( $P=0.9$ ), however if no alarm was sounding, John may still call due to mistaking the phone ring ( $P=0.05$ ). So the strength of each dependency is encoded into the Bayesian Network through the quantification of local conditional probability values.

### 7.3.2. Bayesian Network Features

**Conditional Independence** is key to understanding how evidence propagates through a Bayesian Network. When two events are conditionally independent if they are independent of each other given knowledge of a 3<sup>rd</sup> event. For example in Figure 36(A), the node C is conditionally independent from A given B. ( $P(A|A \cap B) = P(C|B)$ ). For example, if we know the alarm sounded, then the probability that John calls us is independent of what triggered the alarm. For the remaining shapes (Korb & Nicholson 2004):

(A) ( $P(A|A \cap B) = P(C|B)$ ): Conditionally independent

(B) ( $P(A|A \cap B) = P(C|B)$ ): Conditionally independent

(C) ( $P(A|A \cap B) \neq P(C|B)$ ): Conditionally dependent

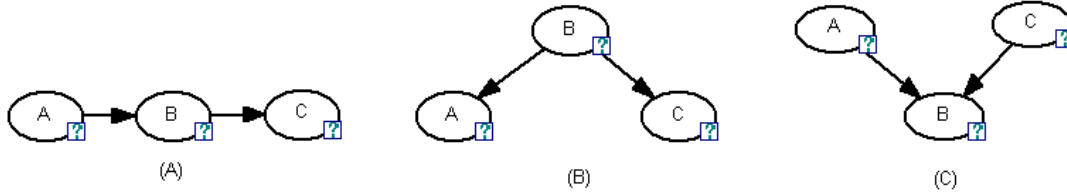


Figure 36: (A) Causal chain (B) common cause (C) common effect

**Joint Probability Distribution.** A Bayesian Network may be considered a graphical representation of a joint probability distribution where:

$$P(x_1, x_2, \dots, x_n) = P(x_1)P(x_2|x_1) \dots P(x_n|x_1, \dots, x_{n-1})$$

For the burglary example, the joint probability distribution for the network can be written as:

$$\begin{aligned} P(\mathbf{X}) &= P(B)P(E|B)P(R|E, B)P(A|R, E, B)P(M|A, R, E, B)P(J|M, A, R, E, B) \\ &= P(B)P(E)P(R|E)P(A|E, B)P(M|A)P(J|A) \end{aligned}$$

Where

$$\mathbf{X} = \{J, M, A, R, E, B\}$$

**Marginal Distributions** can be calculation for a node by integrating out the nuisance parameters from the joint probability distribution. Efficient algorithms exist for conducting this calculation. The marginal distributions for each node in the burglary problem are shown in Figure 37. This represents the probability of each event occurring with no evidence applied to the Bayesian Network model. Note our current belief of a burglary is 0.01.

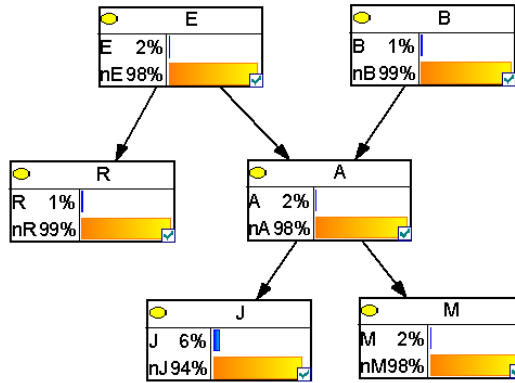


Figure 37: Marginal Distributions for burglary example.

**Diagnostic Reasoning** is where information is known about the symptoms and so our beliefs about the causes are updated. For example if Mary called to say that she heard the alarm, we can apply that event to the Bayesian Network, and our updated update out beliefs that a burglary occurred. Our new belief of burglary with this new evidence is 0.31. (Figure 38 (A))

**Predictive Reasoning** is where information is known about the causes and we update our belief about their effects. For example if we felt an earthquake, we could predict the probability that Mary will call about the alarm. (Figure 38 (B))

**Explaining Away** is where different causes are eliminated to estimate the actual cause. For example, if the alarm was sounded, the three possible causes are earthquake, burglary and malfunction. Despite earthquakes and a burglary having no relationship, by eliminating earthquakes from the causes it increases our belief that a burglary caused the alarm. (Figure 38 (C))

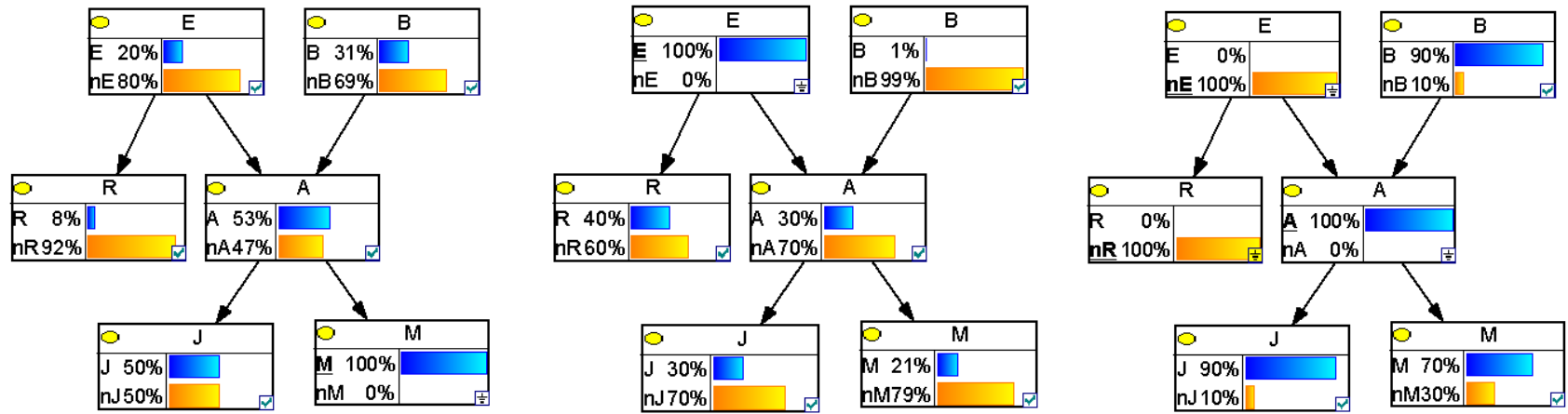


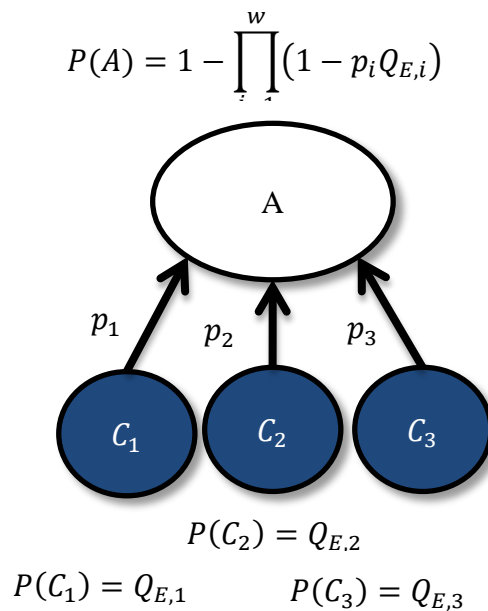
Figure 38: (A) Diagnostic Reasoning (B) Predictive Reasoning (C) Explaining Away

#### 7.4. GDM Bayesian Network Structure

This section will cover how the conceptual GDM covered in section 7.2 is implemented within a Bayesian Network.

##### 7.4.1. *Component Failure Node*

Recall that the component failure may occur due to a failure from any cause, as shown in Figure 39.



**Figure 39: GDM Basic Events**

To represent this in the Bayesian Network, the Conditional Probability Table for the component node, must be calculated given the state of the cause condition nodes. For example, where the system has three possible causes, the probability of component A failing when only one cause condition exists is:

$$P(A|C_1, \overline{C_2}, \overline{C_3}) = p_1$$

$$P(A|\overline{C_1}, C_2, \overline{C_3}) = p_2$$

$$P(A|\overline{C_1}, \overline{C_2}, C_3) = p_3$$

However when multiple causes are present, the probability of component A failing is the union from each failure cause:

$$P(A|C_1, C_2, \overline{C_3}) = 1 - (1 - p_1)(1 - p_2)$$

$$P(A|C_1, \overline{C_2}, C_3) = 1 - (1 - p_1)(1 - p_3)$$

$$P(A|\overline{C_1}, C_2, C_3) = 1 - (1 - p_2)(1 - p_3)$$

$$P(A|C_1, C_2, C_3) = 1 - (1 - p_1)(1 - p_2)(1 - p_3)$$

Therefore the conditional probability table for a component node is:

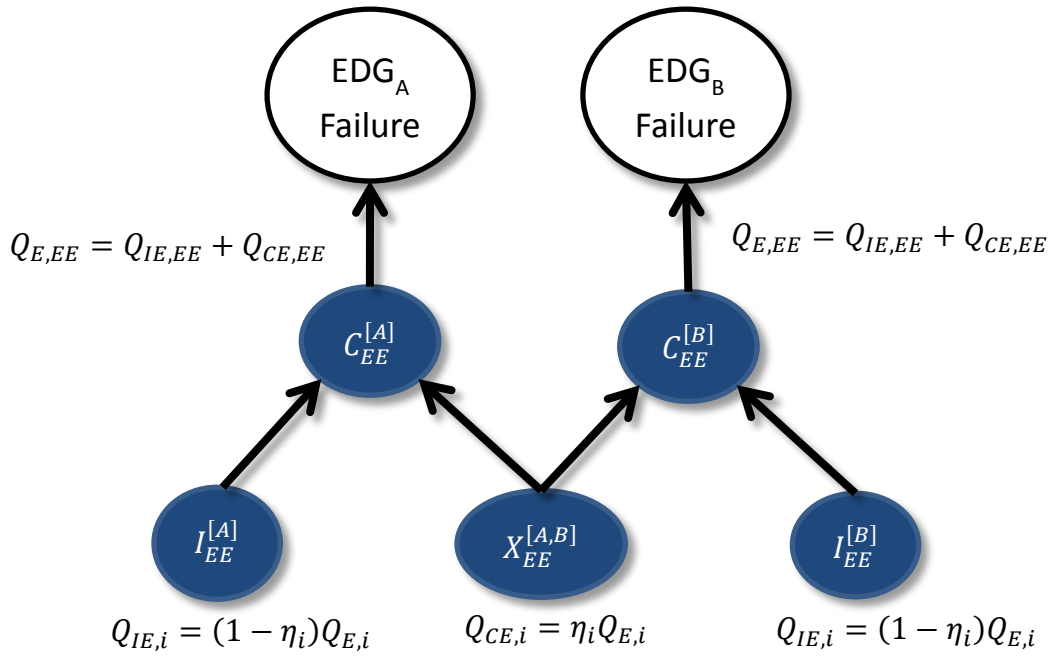
**Table 32: CPT for Control Node**

A State	$C_3$				$\overline{C_3}$			
	$C_2$		$\overline{C_2}$		$C_2$		$\overline{C_2}$	
	$C_1$	$\overline{C_1}$	$C_1$	$\overline{C_1}$	$C_1$	$\overline{C_1}$	$C_1$	$\overline{C_1}$
A	$1 - (1 - p_1)(1 - p_2)(1 - p_3)$	$1 - (1 - p_2)(1 - p_3)$	$1 - (1 - p_1)(1 - p_3)$	$p_3$	$1 - (1 - p_1)(1 - p_2)$	$p_2$	$p_1$	0
$\overline{A}$	$(1 - p_1)(1 - p_2)(1 - p_3)$	$(1 - p_2)(1 - p_3)$	$(1 - p_1)(1 - p_3)$	$1 - p_3$	$(1 - p_1)(1 - p_2)$	$1 - p_2$	$1 - p_1$	1

#### 7.4.2. Cause Condition Nodes

The cause condition nodes required the concept of a local cause condition for each component,  $Q_{E,i}$ , and the ability to propagate that cause condition to other components with probability  $\eta_i$ . This is done by splitting the cause condition probability into a

common cause condition,  $P(X_i) = \eta_i Q_{E,i}$ , and an independent cause condition,  $P(I_i) = (1 - \eta_i)Q_{E,i}$ . See Figure 40 for the conceptual model structure for the cause condition nodes.



**Figure 40: Conceptual cause condition modeling**

The local cause condition,  $C_i$ , is the union of the independent cause condition,  $I_i$ , and common condition,  $X_i$ . These two events are mutually exclusive.

$$P(C_i) = P(I_i) + P(X_i) - P(I_i \cap X_i)$$

where

$$P(I_i \cap X_i) = 0$$

Modeling mutually exclusive events, where only one event has a dependency to other nodes is difficult to achieve in a Bayesian Network model. Therefore the following



modeling options will be discussed:

- Mutually exclusive node state
- Rare event approximation
- Control Nodes

### **Mutually Exclusive Node States**

In a Bayesian Network, mutually exclusive events are modeled through the different node states. Therefore a node may exist which have three states for a cause node:

$$C_i \in \{\text{No Cause, Independent Cause, Common Cause}\}$$

Using such a node to model the cause conditioning within Figure 41 and Figure 42 is problematic for the following reasons:

- Links can only be established from node to node. Therefore there is no way to propagate the common cause condition, without also influencing other nodes with the independent cause condition (Figure 41).
- Bayesian Networks are Directed Acyclic Graphical (DAG) model. This requires the structure of the Bayesian Network to have no cyclic links, and links can only represent one causal direction. Therefore it is not clear whether a link should be established from  $C_{EE}^{[A]} \rightarrow C_{EE}^{[B]}$  or  $C_{EE}^{[B]} \rightarrow C_{EE}^{[A]}$ .(Figure 42)

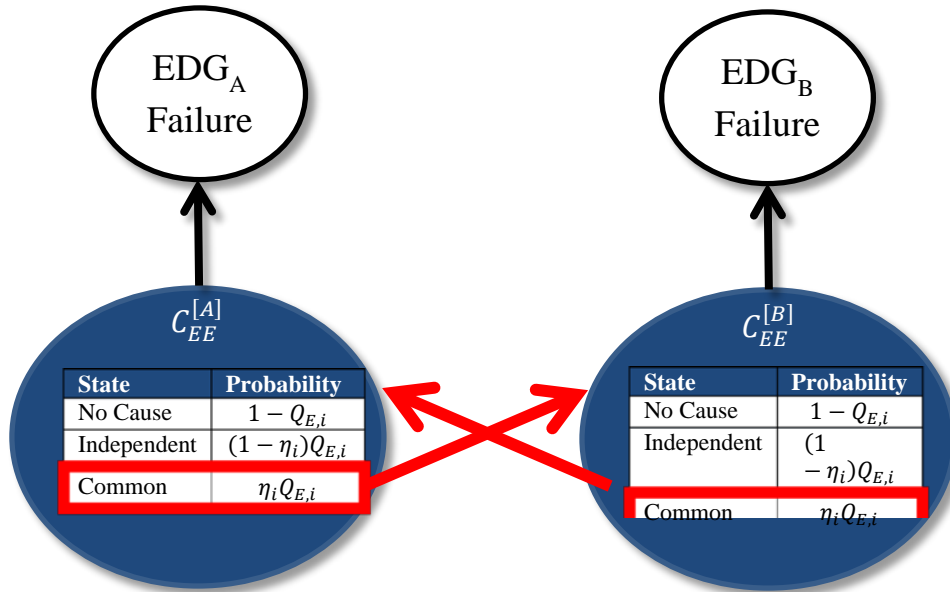


Figure 41: Problem with propagating the common cause condition

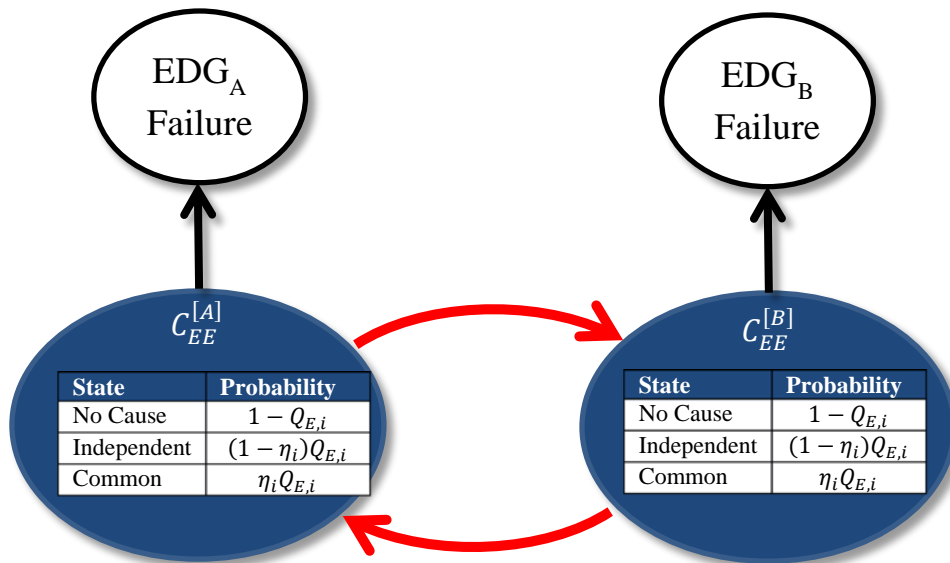


Figure 42: Problem with propagating the common cause condition

It is clear that additional nodes are required to model the independent and common cause condition nodes.

## Rare Event Approximation

Another option is to model the common cause condition and independent cause condition as parent nodes to a local cause node as shown in Figure 65.

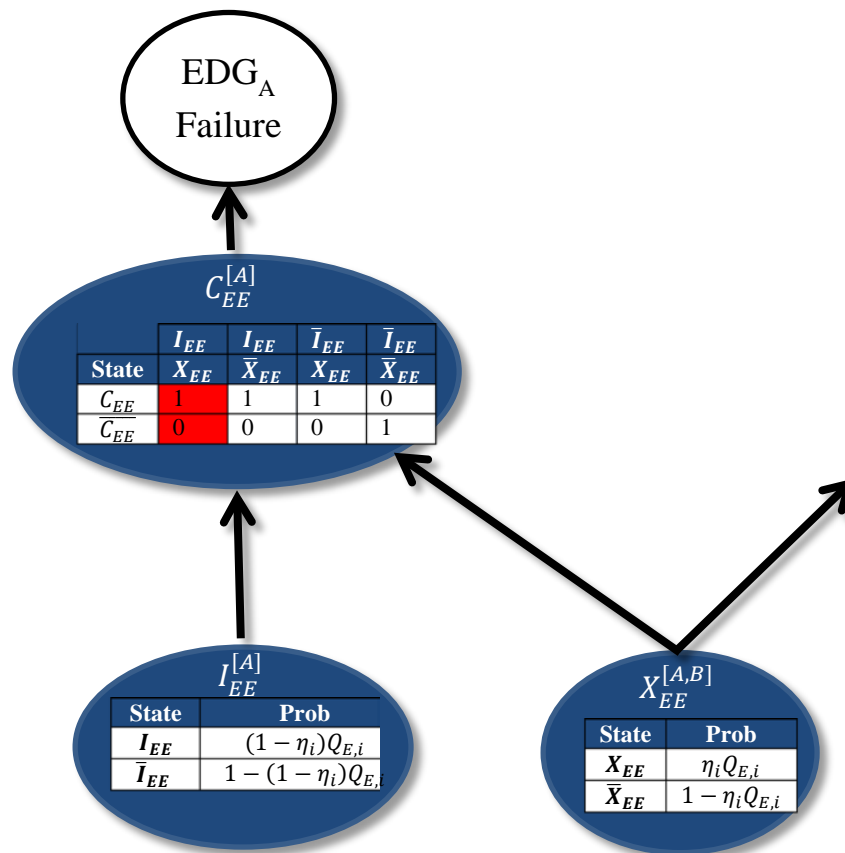


Figure 43: Problem with propagating the common cause condition

During construction of the local cause condition node, the occasion where both the common cause and independent cause condition nodes are true, need to be defined. Although this is an impossible scenario, given the two nodes are mutually exclusive there is little doubt that the result is a cause condition at the local node. The probability of a cause condition at the local node is given as:

$$\begin{aligned}
P(C_i) &= P(C_i|I_i, X_i)P(I_i)P(X_i) + \\
&\quad P(C_i|I_i, \sim X_i)P(I_i)\{1 - P(X_i)\} + \\
&\quad P(C_i|\sim I_i, X_i)\{1 - P(I_i)\}P(X_i) + \\
&\quad P(C_i|\sim I_i, \sim X_i)\{1 - P(I_i)\}\{1 - P(X_i)\}
\end{aligned}$$

$$P(C_i) = P(I_i) + P(X_i) - P(I_i)P(X_i)$$

In order for the events  $I_i$  and  $X_i$  to be treated as mutually exclusive events, the term  $P(I_i)P(X_i)$  is required to be zero. However if  $P(I_i)$  and  $P(X_i)$  are orders of magnitude smaller than 1, the quantity  $P(I_i)P(X_i)$  becomes insignificant. Using a software solution for GDM, this significance could be measured against a threshold, and the simplification made where appropriate.

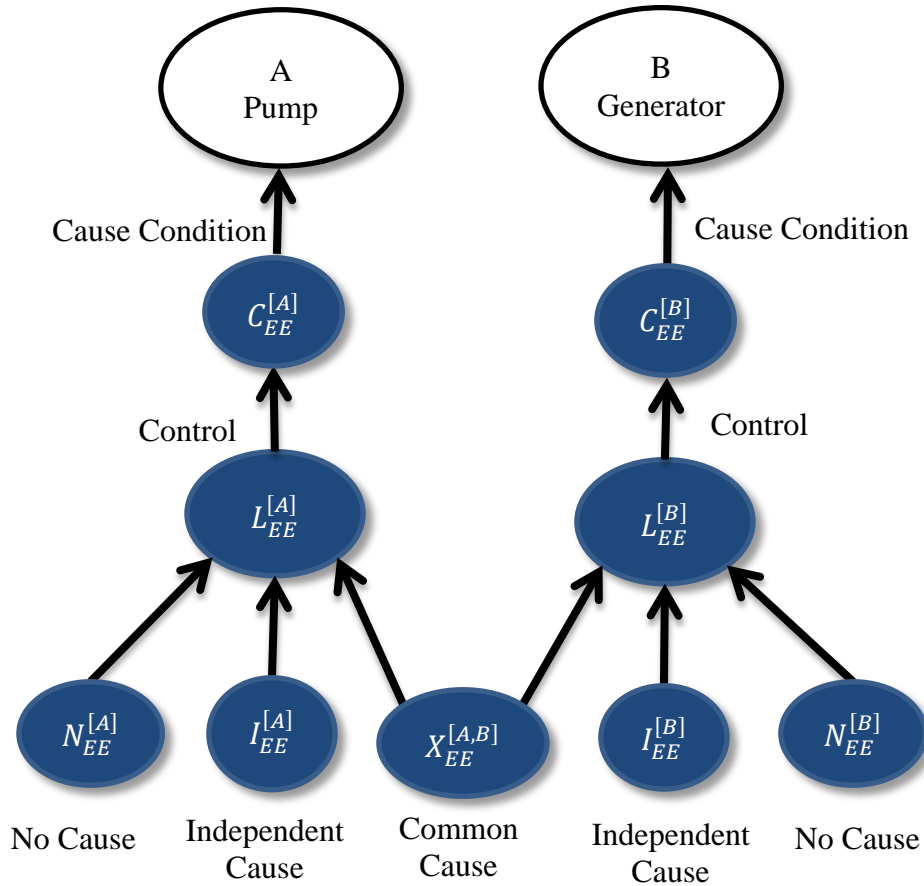
### **Mutually Exclusive Control Nodes**

While CCF events are rare, and CCF events for cause  $i$  even rarer, it may be possible that the cause condition probabilities,  $Q_{IE}$  and  $Q_{CE}$  are not small enough for the rare event approximation to be appropriate. This would occur where the components have low fragility, as seen by the equation,  $Q_{t,i} = p_i Q_{E,i}$ .

Therefore a solution is required to accurately model mutually exclusive events, without assumption. (Fenton et al. 2012) reviewed implementations of mutually exclusive node methods within a Bayesian Network, and proposed the following solution.

All possible mutually exclusive states are created as parent nodes to a control node.

The control node has a child node which represents the desired combined node.



**Figure 44: Structure of mutually exclusive Bayesian network**

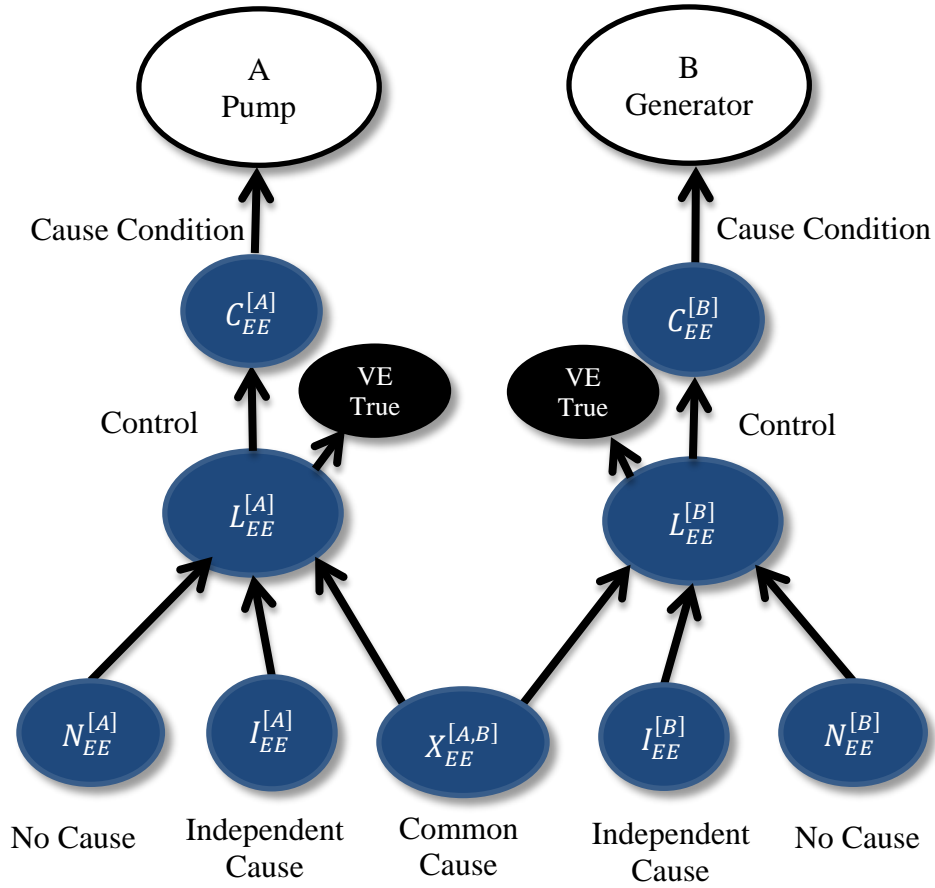
The control node has a state for each mutually exclusive event, plus a “Not Applicable” (NA) state. For each state of the control node, the Conditional Probability Table (CPT) has a ‘1’ where the parental node is true and all other nodes are false. All other combinations of the parental nodes have a ‘1’ in the NA

state. The CPT for the control node is shown in Table 33.

**Table 33: CPT for Control Node**

$L_i$ State	$N_i$				$\bar{N}_i$			
	$I_i$		$\bar{I}_i$		$I_i$		$\bar{I}_i$	
	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$
$N_i$	0	0	0	1	0	0	0	0
$I_i$	0	0	0	0	0	1	0	0
$X_i$	0	0	0	0	0	0	1	0
$NA$	1	1	1	0	1	0	0	1

The control node is forces the probability for each state back to its mutually exclusive state through the application of virtual evidence to the control node. Virtual evidence can be applied directly is many software packages. Virtual evidence can also be applied by creating an additional child node, placing the virtual evidence into the CPT and instantiating the new node true (Pearl 1988, p.46). The structure of the BN model including the virtual evidence node is shown in Figure 45.



**Figure 45: Structure of mutually exclusive Bayesian network with VE**

The weights for the virtual evidence to be applied as virtual evidence is (Fenton et al. 2012):

$$w_{N,i} = \frac{\tau}{[1 - P(I_i)][1 - P(X_i)]} = \frac{\tau}{[1 - (1 - \eta_i)_{E,i}][1 - \eta_i Q_{E,i}]}$$

$$= \frac{Q_{E,i}}{2}$$

$$w_{I,i} = \frac{\tau}{[1 - P(N_i)][1 - P(X_i)]} = \frac{\tau}{Q_{E,i}[1 - \eta_i Q_{E,i}]}$$

$$= \frac{1 - Q_{E,i}(1 - \eta_i)}{2}$$

$$w_{X,i} = \frac{\tau}{[1 - P(N_i)][1 - P(I_i)]} = \frac{\tau}{Q_{E,i}[1 - (1 - \eta_i)Q_{E,i}]}$$

$$= \frac{1 - \eta_i Q_{E,i}}{2}$$

$$w_{NA,i} = 0$$

Where  $\tau$  is a normalizing constant:

$$\tau = \frac{1}{[1 - P(I_i)][1 - P(X_i)]} + \frac{1}{[1 - P(N_i)][1 - P(X_i)]} + \frac{1}{[1 - P(N_i)][1 - P(I_i)]}$$

The CPT for the virtual evidence node is shown in Table 34:

**Table 34: CPT for Virtual Evidence Node**

$VE_i$ State	$N_i$	$I_i$	$X_i$	NA
<b>True</b>	$\frac{Q_{E,i}}{2}$	$\frac{1 - Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 - \eta_i Q_{E,i}}{2}$	0
<b>False</b>	$1 - \frac{Q_{E,i}}{2}$	$\frac{Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 + \eta_i Q_{E,i}}{2}$	1

It can be shown that with the virtual evidence applied, the probability of each state of the control node reflect the mutually exclusive states of the parent nodes. Appendix 3 provide the calculations to show:

$$P(L_i = X_i|V_i) = \eta_i Q_{E,i} = P(X_i)$$



$$P(L_i = I_i|V_i) = Q_{E,i}(1 - \eta_i) = P(I_i)$$

$$P(L_i = N_i|V_i) = (1 - Q_{E,i}) = P(N_i)$$

$$P(L_i = NA|V_i) = 0$$

### Local Cause Condition Node

The control node has four states, however the local cause condition node is only interested in whether there is a cause present or not. Therefore the CPT for the local cause condition node,  $C_i$  is:

**Table 35: CPT for Local Cause Condition Node**

$C_i$ State	$N_i$	$I_i$	$X_i$	$NA$
$C_i$	0	1	1	0
$\bar{C}_i$	1	0	0	1

The marginal probability for the local cause condition node is:

$$\begin{aligned} P(C_i|V_i) &= \sum_j P(C_i | L_{i,j})P(L_{i,j}|V_i) \\ &= (1 - \eta_i)Q_{E,i} + \eta_iQ_{E,i} \\ &= Q_{E,i} \end{aligned}$$

$$\begin{aligned} P(\bar{C}_i|V_i) &= \sum_j P(\bar{C}_i | L_{i,j})P(L_{i,j}|V_i) \\ &= 1 - Q_{E,i} \end{aligned}$$

## Cause Condition Parent Nodes

The Conditional Probability Tables for the three parent nodes,  $X_i, I_i, N_i$  are:

**Table 36: CPT for Common Cause Condition  $X_i$**

$X_i$ State	
$X_i$	$\eta_i Q_{E,i}$
$\bar{X}_i$	$1 - \eta_i Q_{E,i}$

**Table 37: CPT for Independent Cause Condition  $I_i$**

$I_i$ State	
$I_i$	$(1 - \eta_i) Q_{E,i}$
$\bar{I}_i$	$1 - (1 - \eta_i) Q_{E,i}$

**Table 38: CPT for No Cause Condition  $N_i$**

$N_i$ State	
$N_i$	$1 - Q_{E,i}$
$\bar{N}_i$	$Q_{E,i}$

### 7.4.3. Graphical Representation of GDM

This section consolidates the Bayesian Network structure and conditional probability tables, and provides an example structure that would be used for example. An abbreviated graphical representation of GDM will be proposed.

### Complete Representation

Recall that example 1 has the following features as shown in Figure 46 and Table 39.

- Two components
- Three cause types
- Shared coupling factors for each cause

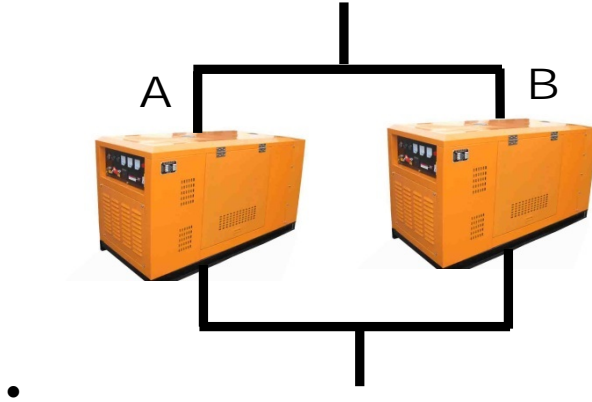


Figure 46: Reliability block diagram for example 1- Two train EDG system

Table 39: Qualitative dependency assessment for example 1

Component	Install Procedure	Maintenance Staff	Location
EDG 1 (A)	EDG IP	Team X	Room Y
EDG 2 (B)	EDG IP	Team X	Room Y

Figure 47 shows the structure of a GDM model for example 1.

**Virtual Evidence Node,  $V_i$**

State	$N_i$	$I_i$	$X_i$	NA
True	$\frac{Q_{E,i}}{2}$	$\frac{1 - Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 - \eta_i Q_{E,i}}{2}$	0
False	$1 - \frac{Q_{E,i}}{2}$	$\frac{Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 + \eta_i Q_{E,i}}{2}$	1

**Control Node,  $L_i$**

State	$N_i$				$\bar{N}_i$			
	$I_i$		$\bar{I}_i$		$I_i$		$\bar{I}_i$	
	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$
$N_i$	0	0	0	1	0	0	0	0
$I_i$	0	0	0	0	0	1	0	0
$X_i$	0	0	0	0	0	0	1	0
NA	1	1	1	0	1	0	0	1

**Component Node, A and B**

State	$C_3$				$\bar{C}_3$			
	$C_2$		$\bar{C}_2$		$C_2$		$\bar{C}_2$	
	$C_1$	$\bar{C}_1$	$C_1$	$\bar{C}_1$	$C_1$	$\bar{C}_1$	$C_1$	$\bar{C}_1$
A	$1 - (1 - p_1)(1 - p_2)(1 - p_3)$	$1 - (1 - p_2)(1 - p_3)$	$1 - (1 - p_1)(1 - p_3)$	$p_3$	$1 - (1 - p_1)(1 - p_2)$	$p_2$	$p_1$	0
$\bar{A}$	$(1 - p_1)(1 - p_2)(1 - p_3)$	$(1 - p_2)(1 - p_3)$	$(1 - p_1)(1 - p_3)$	$1 - p_3$	$(1 - p_1)(1 - p_2)$	$1 - p_2$	$1 - p_1$	1

**Common Cause Condition  $X_i$**

$X_i$ State	
$X_i$	$\eta_i Q_{E,i}$
$\bar{X}_i$	$1 - \eta_i Q_{E,i}$

**Independent Cause Condition  $I_i$**

$I_i$ State	
$I_i$	$(1 - \eta_i)Q_{E,i}$
$\bar{I}_i$	$1 - (1 - \eta_i)Q_{E,i}$

**No Cause Condition  $N_i$**

$N_i$ State	
$N_i$	$1 - Q_{E,i}$
$\bar{N}_i$	$Q_{E,i}$

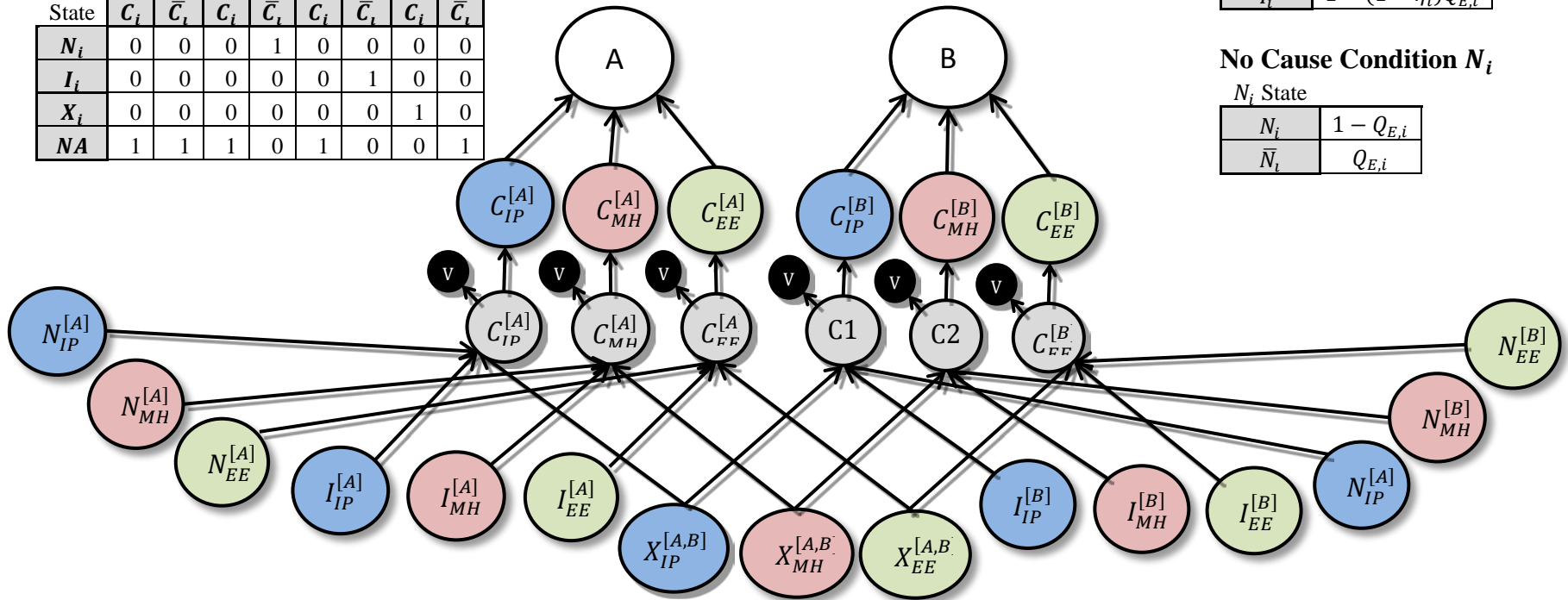


Figure 47: Example GDM Bayesian Network structure for example 1

## Compact Representation

Most of the nodes shown in Figure 47 are created to correctly treat the propagation of the cause condition. When displaying such a model, the items of interest to the analyst are:

- The specific coupling factor which components share. For example “Turbine Building” is the specific coupling factor for a shared location.
- The local cause conditions for each component. The analyst may want to instantiate a known cause condition for event assessment.

Therefore a compact representation of the GDM would be useful. The nodes  $L_i, N_i, I_i, X_i$  can be combined to describe a particular feature of the system which couples components together. This combination of nodes in a fully specified GDM model to a compact representation can be seen in Figure 48 and Figure 49 respectfully. This compact representation can be implemented in Bayesian Network software using Object Orientated Bayesian Network nodes.

The compact representation allows the structure of the Bayesian Network to be easily conveyed to the analyst, while a software solution may decide whether a complete mutually exclusive implementation is adopted or whether the rare event approximation model is used.

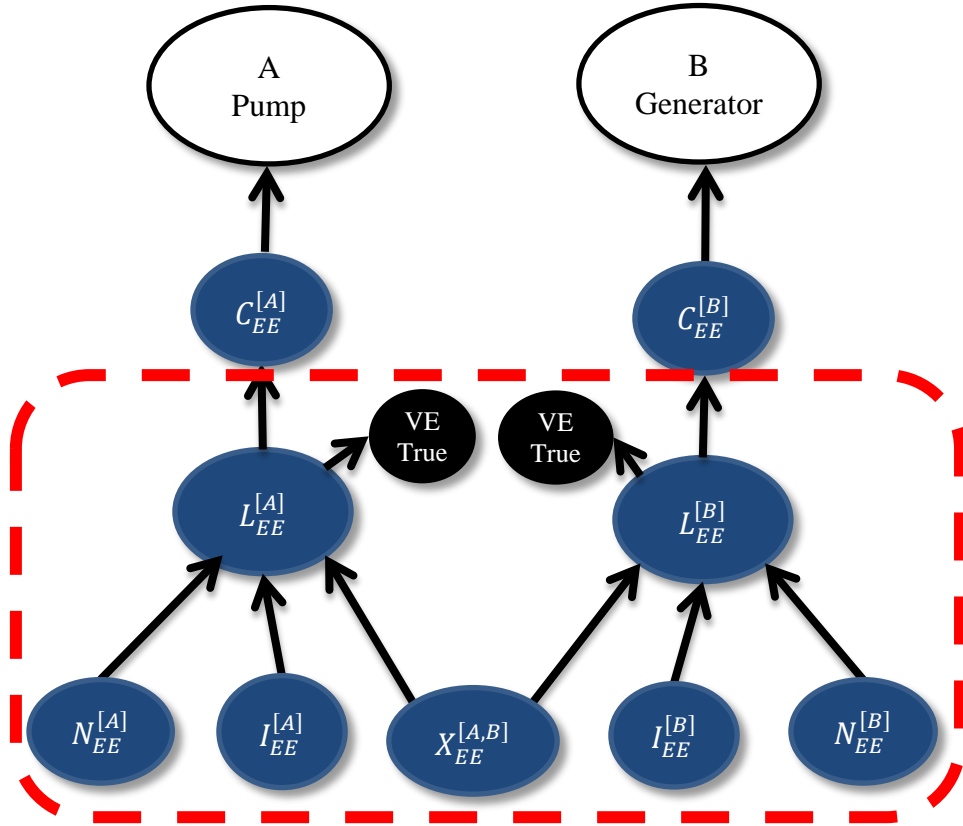


Figure 48: Nodes from the GDM Model which may be combined for visual representation

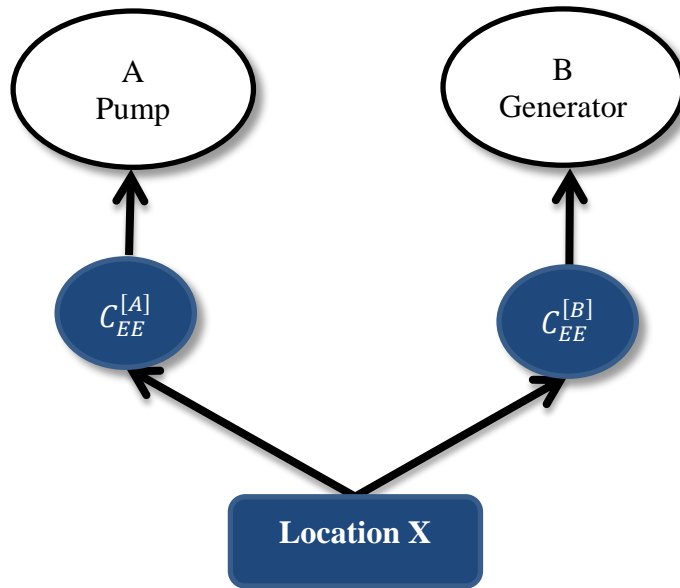


Figure 49: Compact representation of GDM

### 7.5. Parameter Estimation

For each cause the Bayesian Network is fully specified once the three parameters  $p_i, Q_{E,i}, \eta_i$  are known. However, using data from the NRC failure databases, the observable quantities are:

- The failure rate for a component due to cause  $i, Q_{t,i}$ .
- The propensity for common cause failure due to cause  $i$  in a perfectly symmetrical CCCG,  $\alpha_{2,i}$ . The choice of parameter  $\alpha_{2,i}$  will be discussed in section 7.5.2.

A classical and Bayesian formulation will be discussed for each parameter estimate.

#### 7.5.1. *GDM Relationship to $Q_{t,i}$*

The failure probability for cause  $i, Q_{t,i}$  is an observable metric which can assist in the calculation of the GDM parameters through the relationship:

$$Q_{t,i} = p_i Q_{E,i}$$

This section will detail the estimation method for  $Q_{t,i}$  from impact vectors.

## Classical Estimation

The frequentist point estimate for the failure probability of a component due to cause  $i$ ,  $Q_{t,i}$  is:

$$Q_{t,i} = \frac{n_{F,i}}{N_1}$$

$n_{F,i}$  = *the total number of failures due to cause  $i$ .*

$N_1$  = *the total number of demands on a single component.*

Note, this estimate is a rate which can be converted to a mission probability metric as discussed in section 7.2.4.

The quantities,  $n_{F,i}$  and  $N_1$  are component event data, as opposed to CCF event data (see Glossary). For example, if two components are in redundancy, assume that when the system is demanded, both components are demanded. In the first demand, component 1 fails. In the second demand, component 2 fails, in the third demand, no components fail. Then each component was demanded 3 times, making a total of 6 component demands for the system with two failures. The failure probability is  $2/6 = 1/3$ .

This number of failure due to a cause,  $n_{F,i}$  and the number of demands on the components,  $N_1$ , may be available from the target system failure data where a cause has been recorded against each failure.



## Bayesian Estimations

Where a Beta distribution prior is used,  $\pi^0(Q_{t,i}; a_i^0, b_i^0)$ , the parameters for the posterior distribution of the cause failure probability,  $\pi(Q_{t,i}; a_i, b_i)$ , is:

$$a_i = a_i^0 + n_{F,i}$$
$$b_i = b_i^0 + N_1 - n_{F,i}$$

The point estimate for  $Q_{t,i}$  is:

$$\hat{Q}_{t,i} = \frac{a_i}{a_i + b_i}$$

The choice of a prior distribution parameters,  $a_i^0$  and  $b_i^0$ , depends on the availability of data and has been discussed in section 6.5.

### 7.5.2. GDM Relationship to $\alpha_{2,i}$

The CCF data measures the strength of a coupling factor through the frequency of CCF events observed. In the AFM, this is quantitatively measured through the use of alpha factors. The calculation of alpha factors for each cause has already been discussed through the calculation of Partial Alpha Factors (see Chapter 6). Therefore  $\alpha_{2,i}$  it is a convenient measure to use for the GDM.

A key difference between GDM and the PAFM, is the calculation of higher multiplicities of failure. GDM uses a an assumption that each component has a

Bernoulli trial in the presence of a cause condition, and will fail with probability  $p_i$ . Therefore the higher multiplicities of failure are not explicitly modeled, which is similar in concept to the Binomial Failure Rate Model (BFRM).

Furthermore the assumption required to estimate the PAFM parameters required that all components within the CCCG were perfectly symmetrical in design, use and dependencies. The higher the size of a CCCG, the less likely it is this assumption is satisfied. Therefore it is proposed that only the second alpha factor is required. This has the advantage that there is likely to be much more data on CCCGs with two components, than larger groups.

In order to obtain a relationship between  $\alpha_{2,i}$  and GDM, the results from the event assessment of a two train, perfect symmetry, system will be analyzed. Recall from section 6.7.2 that the probability of failure for component  $A_i$ , given knowledge of component B failing due to cause  $i$ , is:

$$P(A_i|B_i) = \alpha_{1,i}^2 Q_t + \alpha_{2,i}$$

Following the event assessment calculations for example 1, it can be seen that the  $\alpha_2$  term is the probability of CCF (calculated from the cut set  $\{X_{AB,i}\}$ ), while the remaining term,  $\alpha_{1,i}^2 Q_t$ , is the normalized probability of independent failure for component  $A_i$ .

The same calculation can be done using GDM. The node  $C_i^{[B]}$  is instantiated as true, and the probability for the second component is calculated,  $P(A_i | C_i^{[B]})$ . The calculations to conduct the algebraic solution for each node are cumbersome, therefore the detailed calculations are contained in appendix 4.

$$P(A_i | C_i^{[B]}) = \frac{p_i Q_{E,i} (\eta_i - 1)^2}{1 - Q_{E,i} - Q_{E,i} (\eta_i - 1)} + p_i \eta_i$$

During this calculation it can be seen that the term  $p_i \eta_i$  is the probability of CCF. The remaining term, is the normalized probability of independent failure for component

Therefore, when the GDM model is calculated for two components with perfect symmetry:

$$p_i \eta_i = \alpha_2$$

The estimators for  $\alpha_{2,i}$  were provided in section 6.4.1, and repeated here for completeness.

### **Classical Estimation**

When the failure cause taxonomy is defined in such a way that each cause could only propagate through one coupling factor (the topic of Chapter 4), the frequentist point estimate for the partial alpha factor is::

$$\alpha_{2,i} = \frac{n_{2,i}}{n_{t,i}}$$

where

$$n_{t,i} = \sum_{k=1}^2 n_{k,i}$$

$\alpha_{2,i}$  = *a partial alpha factor which represents the portion of system failure events which resulted in 2 components failing within a common cause component group of size 2. when there was a potential for failure propagation through coupling factor i where  $i \in \{1,2,3,\dots,w\}$*

$n_{k,i}$  = *the number of failure events/frequency which resulted in k components failing within a common cause component group of size m, ( $1 \leq k \leq 2$ ) of coupling factor i where  $i \in \{1,2,3,\dots,w\}$*

$n_{t,i}$  = *the total number of common cause failure events for coupling factor/cause i where  $i \in \{1,2,3,\dots,w\}$ .*

### Bayesian Estimations

Where a Beta distribution prior is used,  $\pi^0(\alpha_{2,i}; a_{2,i}^0, b_{2,i}^0)$ , the parameters for the posterior distribution of the partial alpha factor,  $\pi(\alpha_{2,i}; a_{2,i}, b_{2,i})$ , is:

$$a_{2,i} = a_{2,i}^0 + n_{2,i}$$

$$b_{2,i} = b_{2,i}^0 + n_{t,i} - n_{2,i}$$

The point estimates for each partial alpha factor can be obtained using:

$$\hat{\alpha}_{2,i} = \frac{a_{2,i}}{a_{2,i} + b_{2,i}}$$

The choice of a prior distribution parameters,  $a_{2,i}^0$  and  $b_{2,i}^0$ , depends on the availability of data and has been discussed in section 6.5.

### 7.5.3. Estimation Using Observed Data

The following relationships have been established:

$$Q_{t,i} = p_i Q_{E,i}$$

$$\alpha_{2,i} = p_i \eta_i$$

Where, the terms,  $Q_{t,i}$  and  $\alpha_{2,i}$  are observable. With three unknowns and two equations.

To complete the estimation of the GDM parameters, one of the three parameters must be estimated through other means. Quantification of the following options will be discussed in section 7.6:

- Direct assessment from data which represents a parameter.
- Using constraints from asymmetrical components (see below).
- Assume  $\eta_i = 1$  as per Binomial Failure Rate Model.
- Estimate from parametric failure model, such as human reliability models for human cause conditions,  $Q_{E,i}$  or load strength interference model for  $p_i$ .
- Engineering assessment.
- Solve using data from higher levels of alpha factors.

When components share a coupling factor. The common cause condition must be equal to all components. Where a component has already been parameterized, the remaining components have a third constraint on their parameters, this allowing the system of

equations to be solved. The third equation is:

$$Q_{CE,i} = \eta_i Q_{E,i}$$

Each type of cause lends itself better to different types of estimation techniques for the third parameter (discussed further in section 7.6):

- *Human Error Cause.* The area of Human Reliability Assessments is rich with literature and parametric models. Therefore human causes may be best suited to parametric modeling to estimate  $Q_{E,i}$  or  $\eta_i$ . Failing this, an engineering assessment of  $\eta_i$  would be the next best option.
- *Procedural Error Cause.* Where components are coupled by the same procedure, it is highly likely that if one component is affected, then all shared components may be affected. Therefore procedural errors may be suitable for the assumption  $\eta_i = 1$  or an expert elicitation estimate of  $\eta_i$ .
- *Environmental Error Cause.* Unlike the other causes, environmental cause conditions may be detectable using sensors. In such cases,  $Q_{E,i}$  may be estimated directly from cause condition data. Where this is not possible, the propagation of environmental causes will change between systems depending on the location and building design housing components. Therefore environmental causes may be suitable for an expert elicitation estimate of  $\eta_i$ .

-

#### ***7.5.4. Parameter Uncertainty Calculation***

The GDM parameters are calculated from observable values, and estimated values which have uncertainty distributions. The parameters,  $Q_{t,i}$  and  $\alpha_{2,i}$  are quantified as Beta distributions. The estimation of the third parameter through means described in section 7.6 is likely to have a Beta distribution.

The multiplication/division of random variables distributed with a Beta distribution does not have a closed form. Therefore the uncertainty distribution of the remaining GDM parameters must be calculated using numerical procedures, such as Monte Carlo simulation.

#### **7.6. Parameter Quantification**

For this section it is assumed that the failure taxonomy allows for a one to one, direct relationship between failure causes and coupling factors. This issue is discussed in Chapter 4.

##### ***7.6.1. Direct Estimates***

It may be possible that sufficient data is available to directly estimate the parameters for the GDM. This is only likely for observable cause conditions (even when failure does not occur) such as for extreme external environment, and so will only be briefly covered here.

Note that the direct estimates provided here have used a mission period metric, however they can equally be defined as a rate, such as per demand or per hour and converted to a mission period as detailed in section 7.2.4.

### **Fragility ( $p_i$ )**

If a cause condition can be measured, then the parameter  $p_i$  may be directly estimated from data.

Let  $n_{E,i}$  be the number of mission periods for which the cause condition existed. Let  $n_{F,i}$  be the number of failures due to the cause condition  $i$ . Then the fragility may be directly estimated using:

$$\hat{p}_i = \frac{n_{F,i}}{n_{E,i}}$$

$n_{F,i}$  = *the total number of failures due to cause  $i$ .*

$n_{E,i}$  = *the number of mission periods for which cause condition  $i$  existed.*

Care must be used in definition  $n_{E,i}$  in the context of the mission period.  $n_{E,i}$  is the number of mission periods for which the cause condition existed. This means that one incident which lasts over many mission periods may contribute to multiple counts of  $n_{E,i}$ . Therefore direct estimates of fragility are best suited to short events relative to the mission period.



### **Cause Condition Probability ( $Q_{E,i}$ )**

If a cause condition can be measured, then the cause condition probability,  $Q_{E,i}$  may be directly estimated from data.

Let  $n_{E,i}$  be the number of mission periods for which the cause condition  $i$  existed and  $n_m$  be the number of missions over the period of data collection. Then the error rate may be directly estimated using:

$$\hat{Q}_{E,i} = \frac{n_{E,i}}{n_m}$$

$n_{E,i} =$  *the number of mission periods for which cause condition  $i$  existed.*

$n_m =$  *the total number of mission periods.*

Direct estimates of the cause condition probability is best suited to short events relative to the mission period.

### **Coupling Factor Strength ( $\eta_i$ )**

If a cause condition can be measured in multiple locations, then the coupling factor strength,  $\eta_i$ , may be directly estimated from data.

Let  $n_{IE,i}$  be the number of mission periods for which the cause condition  $i$  existed locally at a component, without occurring at other components. Let  $n_{CE,i}$  be the number of Mission period for which the cause condition  $i$  existed at multiple components. Then

the coupling factor strength may be directly estimated using:

$$\hat{\eta}_i = \frac{n_{E,i}}{n_m}$$

$n_{E,i}$  = *the number of mission periods for which cause condition i existed.*

$n_m$  = *the total number of mission periods.*

Direct estimate of the coupling factor strength is best suited to short events relative to the mission period.

### 7.6.2. Using Impact Vectors and Causes

Where component specific data exists, the evidence required to calculate  $Q_{t,i}$  and  $\alpha_{2,i}$  can be quantified using the impact vector methodology, where the impact vector size has been adjusted to  $m = 2$ .

Consider that the average impact vector for a CCF event can be represented with the inclusion of the failure cause:

$$\bar{I} = [\bar{F}_0, \bar{F}_1, \bar{F}_2][Cause]$$

Then the sum of average impact vectors for J events for a particular cause is:

$$n_{k,i} = \sum_{j=1}^J \bar{F}_k(j)[Cause]$$

*where Cause=i*

$n_{k,i}$  = *the total number of CCF basic events caused by i involving the failure of k similar components.*

Note,  $n_0$  is not included because a failure cause cannot be determined when there was no failure.

The quantities required to estimate the  $Q_{t,i}$  can now be calculated as:

$$n_{F,i} = \sum_{k=1}^2 kn_{k,i}$$

$$n_t = \sum_{k=1}^2 n_k$$

$$N_1 = 2 \left( n_t + \sum_{j=1}^J \bar{F}_0(j) \right)$$

The quantities required to estimate the  $\alpha_{2,i}$  can now be calculated as:

$$n_{t,i} = \sum_{k=1}^2 n_{k,i}$$

Using impact vectors to quantify these observable parameters allows for the ability to use currently accepted mapping rules from systems with different CCCG sizes and to include uncertainty around observed partial failures, coupling factors and time delays as provided by the impact vector methodology.

Generic data sources in the form of impact vectors may also be used, with an engineering assessment of the weighting factor based on the strength of the data from a particular source. This method is discussed in detail in section 6.5.2.

### ***7.6.3. Coupling Factor Strength Assumption***

One means to solve the simultaneous equations is to assume that the coupling factor strength is equal to one ( $\eta_i = 1$ ). This assumption would be equivalent to the Binomial Failure Rate Model which explicitly models that all components share a shock equally.

Some cause conditions are better suited to this assumption. Candidates include environmental causes where components share the same environment or procedural causes where the same procedure is used on multiple components. Human coupling factors are not well suited to this assumption, as the occurrence of the same human error on multiple components is probabilistic.

### ***7.6.4. Existing Parametric Model Estimate***

An advantage of the GDM parameters is that they can be interpreted according to the physics of CCF, as discussed in section 7.2.4. Many techniques have been developed to estimate these quantities. This section will provide an overview, however the use and integration of such models is left as a recommendation for future research.

Many of the causes related to CCF are related to human error such as installation human error (IH), maintenance human error (MH), operations human error (OH). The GDM parameters which require estimation are (Bell & Holroyd 2009):

- $Q_{E,i}$  the probability of an existing human cause condition which could fail the component. This quantity may be directly related to the Human Error

Probability which is the objective of almost all Human Reliability Assessment (HRA) models such as THERP and SLIM-MAUD.

- $\eta_i$  the probability that a human error may be repeated to other components. This probability is more difficult to estimate using HRA, however it is directly related to the Performance Shaping Factors used to conduct such assessments, and a literature review may find suitable methods are available.
- $p_i$  the fragility of a component in the presence of a human error. This component is generally not considered in HRA, and is usually considered to equal 1, making the HEP the failure probability due to human reliability. A literature review may reveal consideration for component fragility in the presence of human error.

The Human Reliability Assessment models may incorporate data, but also allows for estimation procedures using generic values to adjust estimates through performance shaping factors. Only one GDM parameter requires estimation in order to solve the simultaneous equations if also using available impact vector data.

There are no known parametric modeling techniques for procedural causes. Many environmental models exist for the prediction of environmental conditions (Steppeler et al. 2003), however a detailed literature review and analysis is required to determine their suitability for integration with GDM.

#### **7.6.5. Prior Distributions**

Prior distributions are required for the observable quantities,  $Q_{t,i}$  and  $\alpha_{2,i}$  and most importantly for the estimation of the GDM parameters required to solve the simultaneous equations.

Section 6.5.3 provides detail on the use of the following informative prior distributions.

- *Population Variability Prior.* This method is particularly well suited to the estimation of the component fragility  $p_i$ , as the parameter is a function of the component, which may be reasonably consistent between systems. The parameter  $Q_{E,i}$  is mostly a function of the operating environment and is likely to change between systems. The parameter  $\eta_i$  is mostly a function of the system and support system design, and is better suited to engineering judgment.
- *Expert Elicitation.* While all parameters may have expert elicitation conducted, the parameter  $\eta_i$  is particularly well suited to this method, due to its close relationship with the physical interpretation of a CCF, and due to the need to

customize this estimate based on characteristics of the system. When expert elicitation is used for any GDM parameter, such an estimate can account for specific defenses or vulnerabilities which exist in the target system.

Section 6.5.4 provides detail on the reasons to use a non-informative prior and provides an overview of common Beta distribution non-informative priors that may be used.

### 7.7. *GDM in System Analysis*

The General Dependency Model uses a similar overall process for conducting a CCF analysis as Chapter 2. The following main differences are apparent.

*Identification of CCCGs.* The AFM methodology requires the identification of CCCGs based on a qualitative assessment of similarities between components. The focus is on redundant systems, and on identical components. The GDM methodology requires the same qualitative analysis which will identify common features to any of the components. This assessment can be conducted without regard to the component type or redundancy configuration. However the definition of the common feature is important in defining a coupling factor strength later in the process.

*PRA Structure.* The current AFM methodology modifies the PRA model by splitting basic events into CCBEs. This procedure is manual and obfuscates the original PRA

structure. The GDM methodology does not require the modification of the basic event node structure. Instead the dependencies are modeled through the Bayesian Network which links to the original PRA basic events. This however is not feasible as a manual task, and so software is required to automate this. The advantage with this approach is that the detail of dependency modeling can be hidden from the PRA user, except for the details required to conduct the analysis (see section 7.4.3 for a description of the compact graphical representation of the GDM).

*Parameter Quantification.* The GDM model has parameters for each failure cause which results in many more parameters requiring estimation. Furthermore the GDM parameter estimation procedure described in section 7.5 and 7.6 shows that the parameter estimates are not directly observable and may require further modeling or non-data informed methods to complete parameter quantification. The AFM procedure allows for the estimation of parameters directly from the observable data, without solving for extra parameters.

This section describes the system analysis procedure when using the GDM. The two examples used to demonstrate this procedure are the same two examples described in Chapter 2 and Chapter 6. A two train EDG system with perfect symmetry and a system with two EDG and three pumps with mixed redundancy.

Unlike Chapter 2 and 6, the final quantification of the model will be conducted using



Bayesian Network software, instead of manual calculations.

### 7.7.1. Qualitative Analysis

The purpose of the qualitative analysis is to identify common features between components. The analyst is required to define a threshold as to how far they would consider components to share the same feature. This definition of the boundaries for a common characteristic should be documented. The barriers which prevent failure propagation within this common characteristic will be considered a defense and quantified by  $\eta_i$ . For example, a common location may be defined as being in the same building. Where components are in separate rooms, this separation will be recognized through a  $\eta_i < 1$ . Note that it is possible to model multiple levels of influencing factors and dependency, which will be discussed in section 7.11.1.

The qualitative assessment for example 1 is shown in Table 40.

**Table 40: Qualitative dependency assessment for example 1**

<b>Component</b>	<b>Install Procedure</b>	<b>Maintenance Staff</b>	<b>Location</b>
EDG 1 (A)	EDG IP	Team X	Room Y
EDG 2 (B)	EDG IP	Team X	Room Y

*Installation Procedure (EDG IP)*. Is the same EDG installation procedure for both components.

*Maintenance Staff (Team X)*. Defined as common at the team level. The team members

change, and a different person may be conducting the maintenance on each EDG.

*Location (Room Y).* The two EDGs are next to each other in the same room.

The qualitative assessment for example 2 is shown in Table 41.

**Table 41: Qualitative dependency assessment for example 2**

<b>Component</b>	<b>Install Procedure</b>	<b>Maintenance Staff</b>	<b>Location</b>
EDG 1 (E1)	EDG	Team X	Room Y
EDG 2 (E2)	EDG	Team X	Room Y
Pump 1 (P1)	Pump V1.1	Team X	Room Y
Pump 2 (P2)	Pump V2.8	Team X	Room Y
Pump 3 (P3)	Pump V1.1	Team Y	Room X

*Installation Procedure:*

- (EDG) Is the same EDG installation procedure for both EGD components.
- (Pump V1.1) Installation procedure for pump at plant commissioning.
- (Pump V2.8) Installation procedure from a different company, installed at a later date.

*Maintenance Staff:*

- (*Team X*). Defined as common at the team level. Team X is the onsite maintenance team. The team members change, and a different person may be conducting the maintenance on different components.

- (*Team Y*). Defined at the team level. Team Y is an offsite contractor. The people provided for this task may change, but are always from the same company.

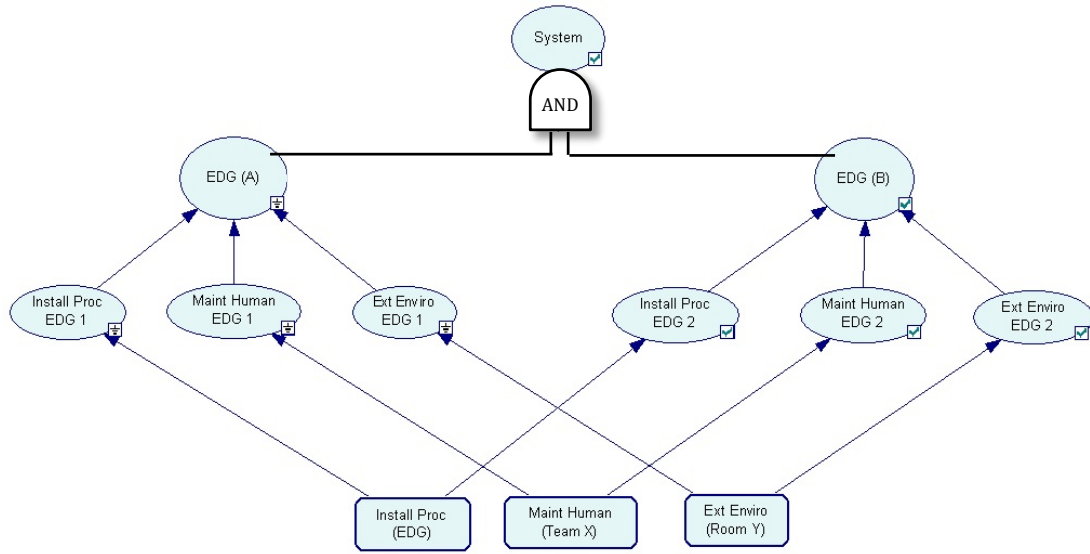
*Location.*

- (Room X) Defined at the room level. Each room has separate climate control.
- (Room Y) Defined at the room level. Each room has separate climate control.

**7.7.2. Create GDM Structure with Common Characteristics**

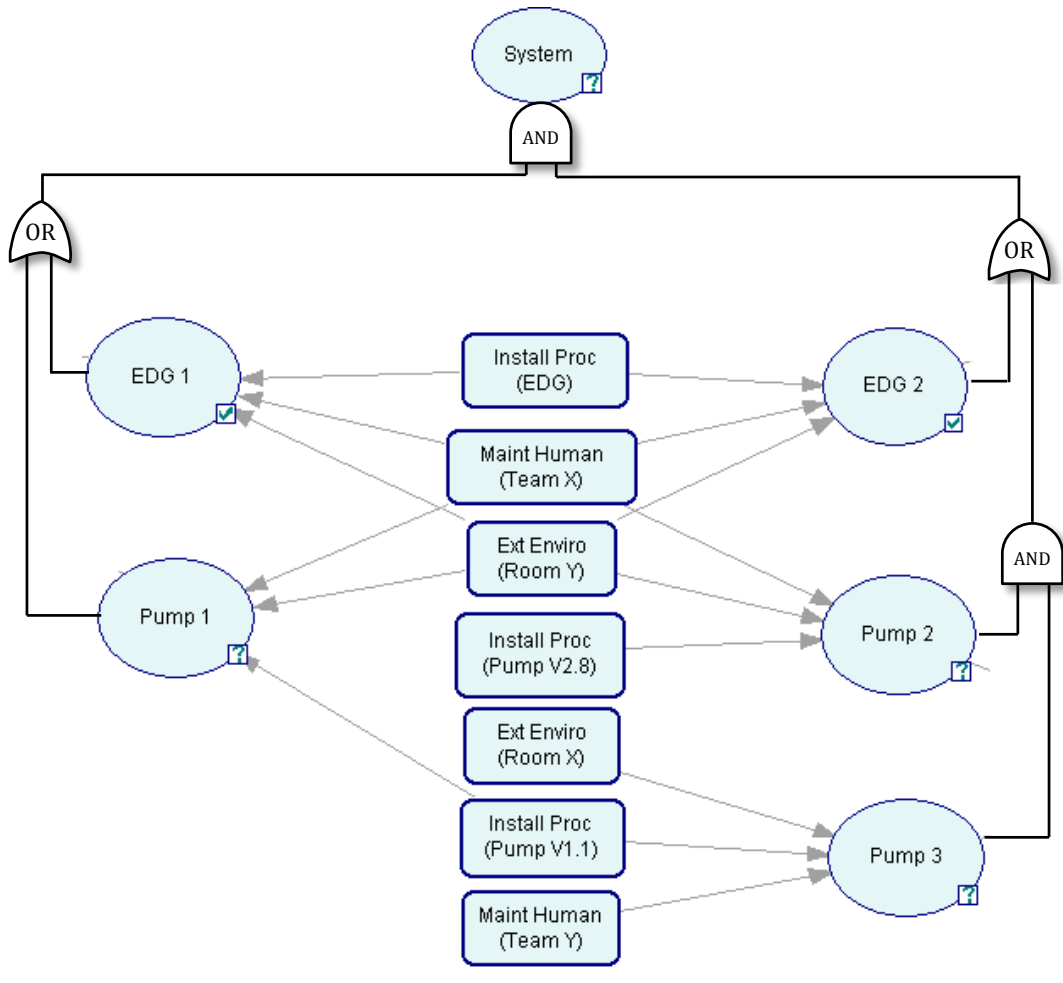
Nodes are created representing the shared characteristics of the system. The dependencies between the component and the shared characteristics are created with links.

Figure 50 shows the GDM structure for example 1. Note that the local conditions are linked where the qualitative analyses showed they shared a condition. Furthermore the shared conditions are named, not just by the type of cause, but by the specific shared characteristic (eg. Room Y).



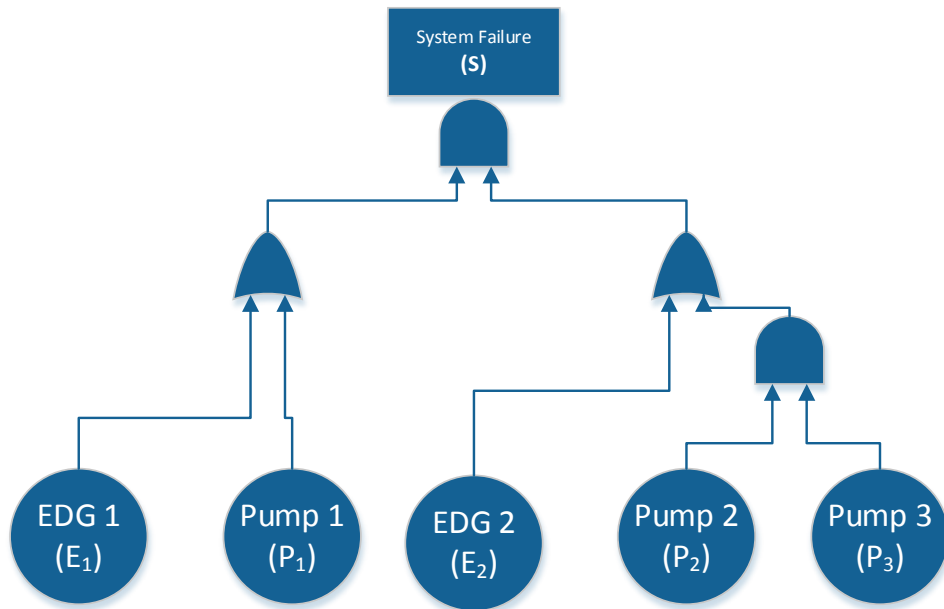
**Figure 50: GDM Structure for Example 1**

Figure 51 shows the GDM structure for example 2. The local cause conditions have been removed for brevity. Again, nodes represent the shared characteristics with links to create the dependency between the local conditions of each component.



**Figure 51: GDM Structure for Example 2**

While Figure 51 has been arranged for presentation in this thesis, in a software application it may be presented to the user using ‘properties’ of the basic event as shown in Figure 52.



Component	EDG	Pump	EDG	Pump	Pump
Location	Room Y	Room Y	Room Y	Room Y	Room X
Install Proc	EDG	Pump1.1	EDG	Pump2.8	Pump1.1
Maint Team	Team X	Team X	Team X	Team X	Team Y
.....	.....	.....	.....	.....	

**Figure 52: GDM Structure for Example 2 analyst interface**

An alternative representation could be shown in a software implemented solution to this step.

### 7.7.3. Identify Constraints from Observable Quantities

For each cause the parameters  $Q_{E,i}$ ,  $\eta_i$ ,  $p_i$  is required to be estimated. This will be done in two steps:

- Identify constraints from observable quantities.
- Estimate parameters within constraints.

For each cause, the quantities  $Q_{t,i}$  and  $\alpha_{2,i}$  can be calculated as constraints.

For example 1, recall the following impact vectors for the EDG, by cause:

$$n_{\Omega,IP}^{[E]} = [0, \quad 172.2, \quad 2.8]$$

$$n_{\Omega,MH}^{[E]} = [0, \quad 154.35, \quad 3.15]$$

$$n_{\Omega,EE}^{[E]} = [0, \quad 16.45, \quad 1.05]$$

$$n_0^{[E]} = 29400$$

The  $Q_{t,i}$  for each cause can be calculated by calculating the total number of mission periods and the total number of failures for each cause. The  $Q_{t,i}$  estimate and  $\alpha_{2,i}$  estimates are shown in Table 42. Note that  $\alpha_{2,i}$  estimates were calculated in section 6.6.7.

$$N_1 = 2 \left( n_t + \sum_{j=1}^J \bar{F}_0(j) \right)$$

$$= 2(29400 + 175 + 157.5 + 17.5)$$

$$= 59500$$

**Table 42: Estimation of  $Q_{t,i}$  for EDG**

Number of failures	$Q_{E,i}$ Estimate	$\alpha_{2,i}$ Estimate
$n_{F,i} = \sum_{k=1}^2 kn_{k,i}$	$Q_{t,i} = \frac{n_{F,i}}{N_1}$	See section 6.6.7
$n_{F,IP}^{[E]} = 172.2 + (2)(2.8)$ $= 177.80$	$Q_{t,IP}^{[E]} = 2.988e-3$	$\alpha_{2,IP}^{[E]} = 0.016$
$n_{F,MH}^{[E]} = 154.35 + (2)(3.15)$ $= 160.65$	$Q_{t,MH}^{[E]} = 2.700e-3$	$\alpha_{2,MH}^{[E]} = 0.020$
$n_{F,EE}^{[E]} = 16.45 + (2)(1.05)$ $= 18.55$	$Q_{t,EE}^{[E]} = 3.118e-4$	$\alpha_{2,EE}^{[E]} = 0.060$

For example 2, recall the following impact vectors for the Pump, by cause:

$$n_{\Omega,IP}^{[P]} = [0, \quad 26.06625, \quad 0.18375]$$

$$n_{\Omega,MH}^{[P]} = [0, \quad 59.4125, \quad 1.8375]$$

$$n_{\Omega,EE}^{[P]} = [0, \quad 82.52125, \quad 4.97875]$$

$$n_0^{[P]} = 44433$$



The  $Q_{t,i}$  for each cause can be calculated by calculating the total number of mission periods and the total number of failures for each cause. The  $Q_{t,i}$  estimate and  $\alpha_{2,i}$  estimates are shown in Table 43. Note that  $\alpha_{2,i}$  estimates were calculated in section 6.6.7.

$$\begin{aligned}
 N_1 &= 2 \left( n_t + \sum_{j=1}^J \bar{F}_0(j) \right) \\
 &= 2(44433 + 26.25 + 61.25 + 87.5) \\
 &= 89216
 \end{aligned}$$

**Table 43: Estimation of  $Q_{t,i}$  for Pump**

Number of failures	$Q_{E,i}$ Estimate	$\alpha_{2,i}$ Estimate
$n_{F,i} = \sum_{k=1}^2 kn_{k,i}$	$Q_{t,i} = \frac{n_{F,i}}{N_1}$	See section 6.6.7
$n_{F,IP}^{[P]} = 26.42$	$Q_{t,IP}^{[P]} = 2.963e-4$	$\alpha_{2,IP}^{[P]} = 0.007$
$n_{F,MH}^{[P]} = 63.09$	$Q_{t,MH}^{[P]} = 7.071e-4$	$\alpha_{2,MH}^{[P]} = 0.03$
$n_{F,EE}^{[P]} = 92.48$	$Q_{t,EE}^{[P]} = 1.0366e-3$	$\alpha_{2,EE}^{[P]} = 0.0569$

#### **7.7.4. Estimate Parameters within Constraints.**

In the previous section the observable quantities,  $Q_{t,i}$  and  $\alpha_{2,i}$  were calculated. These can now be used as constraints to solve the simultaneous equations:

$$Q_{t,i} = p_i Q_{E,i}$$

$$\alpha_{2,i} = p_i \eta_i$$

$$Q_{CE,i} = \eta_i Q_{E,i}$$

Where,  $Q_{CE,i}$ , is unknown, one of the parameters,  $p_i$ ,  $Q_{E,i}$ ,  $\eta_i$  must be estimated by the means discussed in section 7.6. During this step the analyst must focus on estimating the parameters for each component, not the component group. This is because each component has unique features within the system, as will become evident in example 2.

The GDM parameters will be estimated for example 1. Due to the components being symmetrical, the estimates can be completed for a generic EDG and used for both EDGs. First, the coupling factor strength parameter will be estimated using the common characteristic description provided in section 7.7.1

*Installation Procedure (EDG IP).* The installation procedure is the same for all EDG.

If there was an error on one EDG the same error will be evident on the second EDG.

Therefore the coupling factor strength is estimated as 1.  $\hat{\eta}_{EE}^{[Rm Y]} = 1$ .

*Maintenance Staff (Team X).* Different people conduct the maintenance, however they are from the same team. The probability of an error propagating is low. The coupling

factor estimate is low.  $\hat{\eta}_{EE}^{[Tm X]} = 0.2$ .

*Location (Room Y)*. The two EDGs are next to each other in the same room. If one EDG is subjected to an extreme environment, the other is highly likely to be subjected to the same condition. Therefore the coupling factor strength is estimated as 1.  $\hat{\eta}_{EE}^{[Rm Y]} = 0.8$ .

Using the provided estimates, the remaining parameters may be calculated. The results are shown in Table 44.

**Table 44: GDM Parameter Estimates for example 1 EDG**

Cause	$Q_{t,i}$	$\alpha_{2,i}$	$Q_{E,i}$	$p_i$	$\eta_i$
<b>EDG 1</b>					
IP: EDG	0.002988	0.016000	0.186765	0.016000	1
MH: Team X	0.002700	0.020000	0.027000	0.100000	0.2
EE: Room Y	0.000312	0.060000	0.004157	0.075000	0.8
<b>EDG 2</b>					
IP: EDG	0.002988	0.016000	0.186765	0.016000	1
MH: Team X	0.002700	0.020000	0.027000	0.100000	0.2
EE: Room Y	0.000312	0.060000	0.004157	0.075000	0.8

For example 2, the parameters for the EDG have already been calculated from example 1, therefore only the parameters for the pumps are required to be estimated. As stated at the start of this section, the aim is to quantify the GDM parameters for each component. The pumps within example 2 are not symmetrical and therefore greater care must be used when estimating the parameters.

In order to maintain consistency in the modeling parameters, the common cause

condition failure probability  $Q_{CE,i}$  must be equal for all components which share a characteristic. A review of the EDG data estimates and coupling factor reveals that  $Q_{CE,i}$  has already been estimated for causes MH-Team X and EE-Room Y which is relevant to pump 1 and 2. If  $Q_{CE,i}$  is known the system of equations becomes:

$$Q_{t,i} = p_i Q_{E,i}$$

$$\alpha_{2,i} = p_i \eta_i$$

$$Q_{CE,i} = \eta_i Q_{E,i}$$

Solving for the GDM parameters gives:

$$Q_{E,i} = \sqrt{\frac{Q_{t,i} Q_{CE,i}}{\alpha_{2,i}}}$$

$$\eta_i = \sqrt{\frac{\alpha_{2,i} Q_{CE,i}}{Q_{t,i}}}$$

$$p_i = \sqrt{\frac{\alpha_{2,i} Q_{t,i}}{Q_{CE,i}}}$$

For example, the common cause condition for a maintenance human error from team X is calculated as:

$$\begin{aligned} Q_{CE,MH}^{[Tm X]} &= \eta_{MH}^{[E1]} Q_{MH}^{[E1]} \\ &= (0.2)(0.027) \\ &= 5.4e-3 \end{aligned}$$

Using this quantity and the observable parameters for a pump

( $Q_{t,MH}^{[P]} = 7.071e-4$ ,  $\alpha_{2,MH}^{[P]} = 0.03$ ), the GDM parameters for pump 1 and 2, cause MH can be solved:

$$Q_{E,MH}^{[P,Tm X]} = \sqrt{\frac{Q_{t,i} Q_{CE,i}}{\alpha_{2,i}}} = 1.128e-2$$

$$\eta_{MH}^{[P,Tm X]} = \sqrt{\frac{\alpha_{2,i} Q_{CE,i}}{Q_{t,i}}} = 0.479$$

$$p_{MH}^{[P,Tm X]} = \sqrt{\frac{\alpha_{2,i} Q_{t,i}}{Q_{CE,i}}} = 6.268e-2$$

**Table 45** Table 45 shows the parameter estimates using the constraints and engineering assessment. Numbers in red denote where the parameter estimate require  $Q_{CE,i}$  as a constraint in asymmetrical modeling.

With the pump parameters MH: Team X and EE:Room Y being estimated using  $Q_{CE,i}$  the remaining parameters require an engineering assessment to quantify at least one parameter.

*Installation Procedure (Pump IPV1.1)*. The installation procedure is the same for pump 1 and 3. If there was an error on one pump the same error will be evident on the second EDG. Therefore the coupling factor strength is estimated as 1.  $\hat{\eta}_{IP}^{[P1.1]} = 1$ .

*Installation Procedure (Pump IPV2.8).* Pump 2 does not share an installation procedure with another component; therefore the definition of common condition cannot be quantified.  $\hat{\eta}_{IP}^{[P2.8]} = 1$ .

*Maintenance Staff (Team Y).* Pump 3 does not share a maintenance team with another component; therefore the definition of common condition cannot be quantified.  $\hat{\eta}_{MH}^{[Tm Y]} = 1$ .

*Location (Room X).* Pump 3 does not share a room with another component; therefore the definition of common condition cannot be quantified.  $\hat{\eta}_{EE}^{[Tm X]} = 1$ .

**Table 45: GDM Parameter Estimates for example 2**

Cause	$Q_{CE,i}$	$Q_{t,i}$	$\alpha_{2,i}$	$Q_{E,i}$	$p_i$	$\eta_i$
<b>EDG 1</b>						
IP: EDG	0.18676	0.00299	0.01600	0.18676	0.01600	1.000
MH: Team X	0.00540	0.00270	0.02000	0.02700	0.10000	0.200
EE: Room Y	0.00333	0.00031	0.06000	0.00416	0.07500	0.800
<b>EDG 2</b>						
IP: EDG	0.18676	0.00299	0.01600	0.18676	0.01600	1.000
MH: Team X	0.00540	0.00270	0.02000	0.02700	0.10000	0.200
EE: Room Y	0.00333	0.00031	0.06000	0.00416	0.07500	0.800
<b>Pump 1</b>						
IP: Pump 1.1	0.04233	0.00030	0.00700	0.04233	0.00700	1.000
MH: Team X	0.00540	0.00071	0.03000	0.01128	0.06268	0.479
EE: Room Y	0.00333	0.00104	0.05690	0.00778	0.13318	0.427
<b>Pump 2</b>						
IP: Pump 2.8	0.04233	0.00030	0.00700	0.04233	0.00700	1.000
MH: Team X	0.00540	0.00071	0.03000	0.01128	0.06268	0.479
EE: Room Y	0.00333	0.00104	0.05690	0.00778	0.13318	0.427
<b>Pump 3</b>						

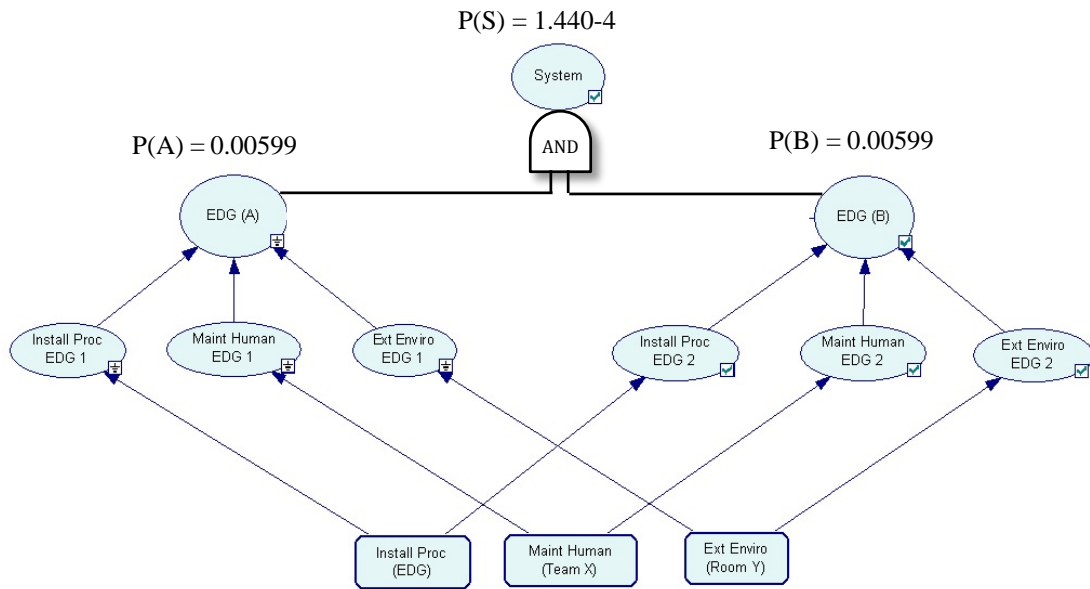
IP: Pump 1.1	0.04233	0.00030	0.00007	0.04233	0.00700	1.000
MH: Team Y	0.02357	0.00071	0.00054	0.02357	0.03000	1.000
EE: Room X	0.01822	0.00104	0.00039	0.01822	0.05690	1.000

### 7.7.5. Calculate Model

Once the GDM parameters have been estimated, the Conditional Probability Tables for the Bayesian Network can be populated according to section 7.4. The model can then be calculated using software.

For example 1 the completed Bayesian Network is shown in Figure 53. The probability of system failure is  $1.440e-4$  which is close to the same calculation using the AFM and PAFM of  $1.546e-4$ . The GDM estimate is lower because the failure rate from each cause is being treated as an independent event, instead of a simple sum. This causes each component reliability to be lower than the point estimate from the combined data. When component failures rates are not rare events (greater than 0.01), the difference between AFM and GDM system estimates will become more pronounced.

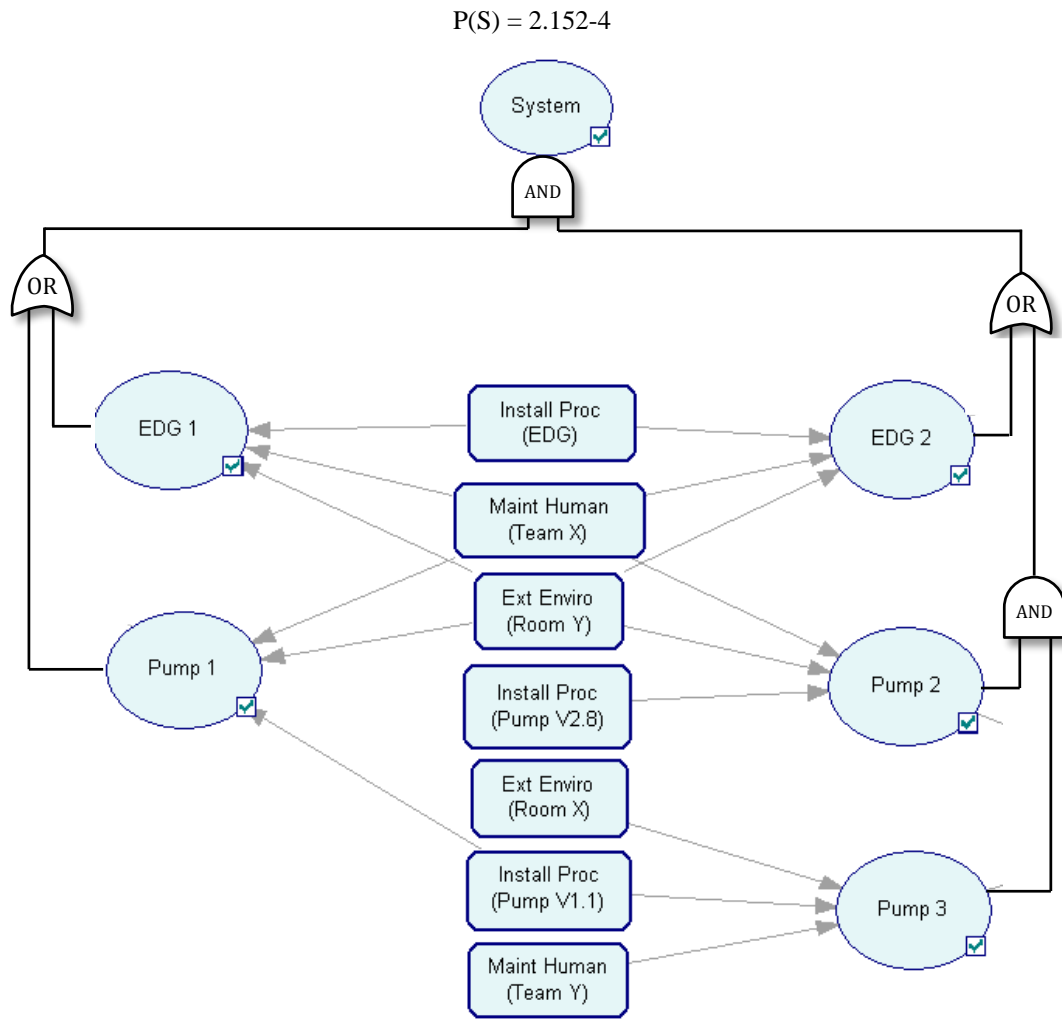
Furthermore, for larger CCCGs the AFM and the GDM model will not be equivalent, as the higher alpha factors have not been used to calculate the GDM parameter estimations.



**Figure 53: GDM system analysis results for example 1**

The completed Bayesian Network for example 2 is shown in Figure 54. The system failure probability is  $2.152e-4$  which is higher than the AFM estimate and PAFM of  $1.668e-4$ . This is due to the additional dependencies between the pumps and EDG which have been modeled.





**Figure 54: GDM system analysis results for example 2**

### 7.8. GDM in Event Assessment

A strength of the GDM model is the flexibility when conducting event assessments.

The three event assessment scenarios will be presented:

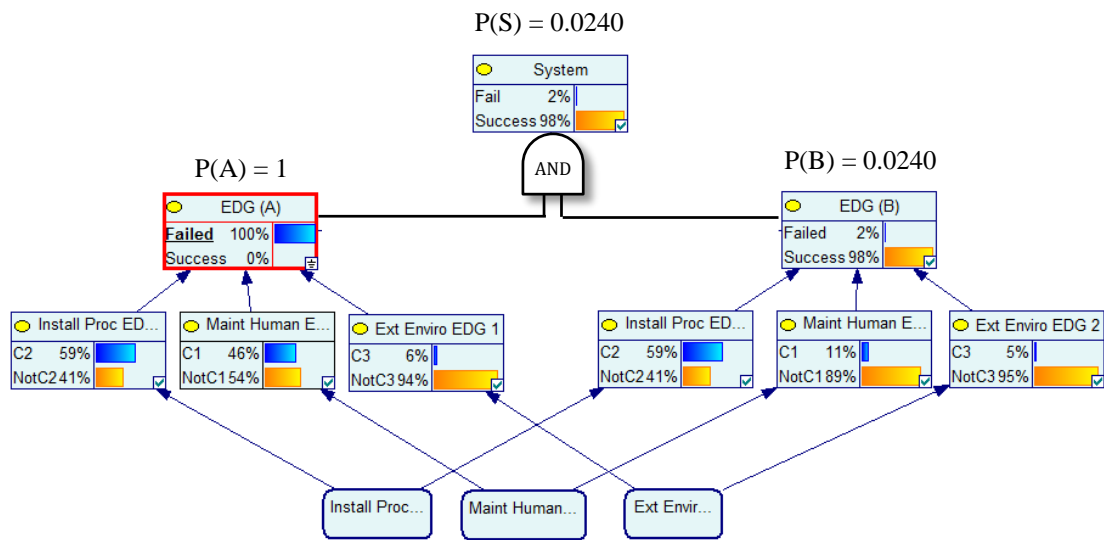
- Event assessment with knowledge of a component failure

- Event assessment with knowledge of a component failure and failure cause
- Event assessment with virtual evidence about the component failure cause

### 7.8.1. Knowledge of Failure

The procedure for conducting event assessment using the Bayesian Network is very straight forward. The analyst applies evidence to the node in the software package and the other node values are updated.

Therefore the probability of system failure given B has failed is shown in Figure 55.



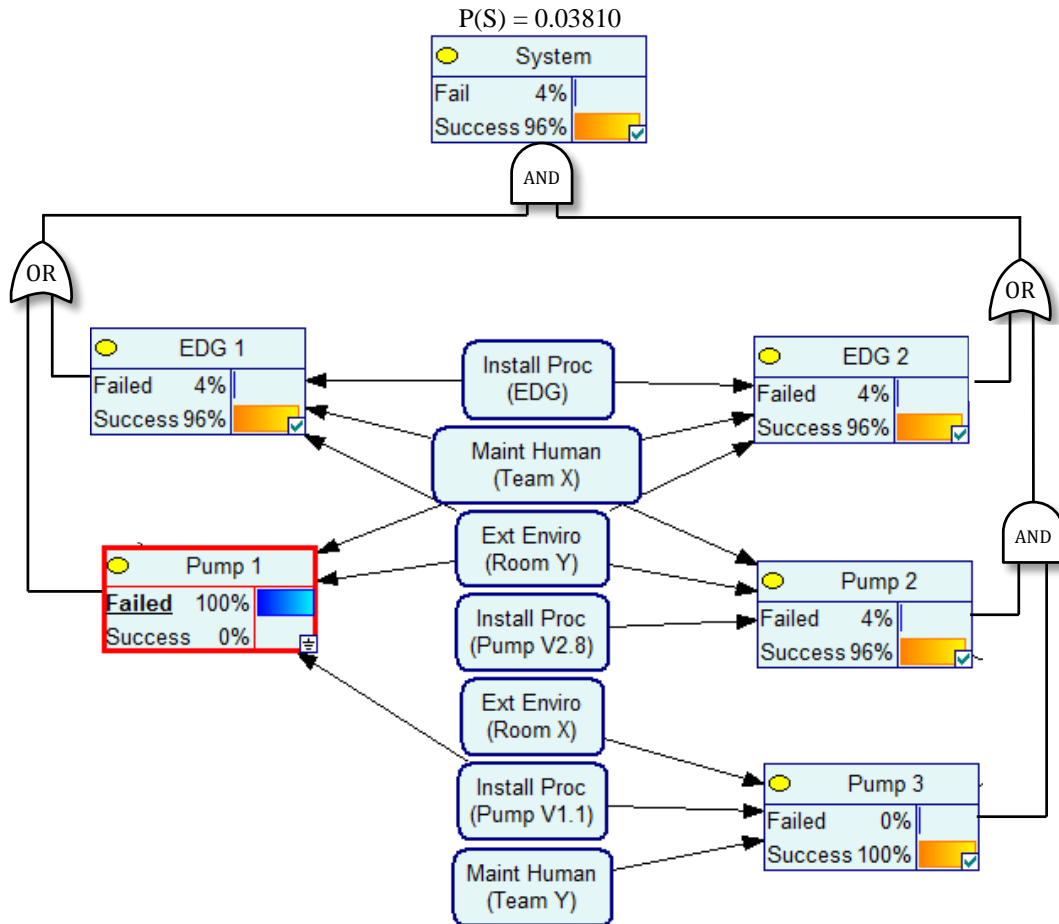
**Figure 55: Event assessment for component A failing using GDM**

The event assessment using GDM gives  $P(S|B) = 0.0240$ , which is slightly less than the AFM and PAFM estimate of 0.0258. This is due to GDM modeling failures from each cause as independent events, as opposed to the PAFM which mostly treats them as mutually exclusive. Example 1 shows that GDM may be considered equivalent for

a two train symmetrical system to the AFM.

GDM also presents a diagnostic function by estimating which local cause is likely to have caused the failure. As can be seen in Figure 55, the most likely cause is an installation error at 59%. Furthermore the probability of a local cause condition existing at component B has also increased. For example the probability that a maintenance human cause condition is at component B increased from 0.03 to 0.11. This information may be useful for diagnostics and guiding root cause analysis. This demonstrates the visual advantages to modeling CCF within a Bayesian Network.

The event assessment of P1 failing on example 2 is shown in Figure 56.



**Figure 56: Event assessment for component P1 failing using GDM**

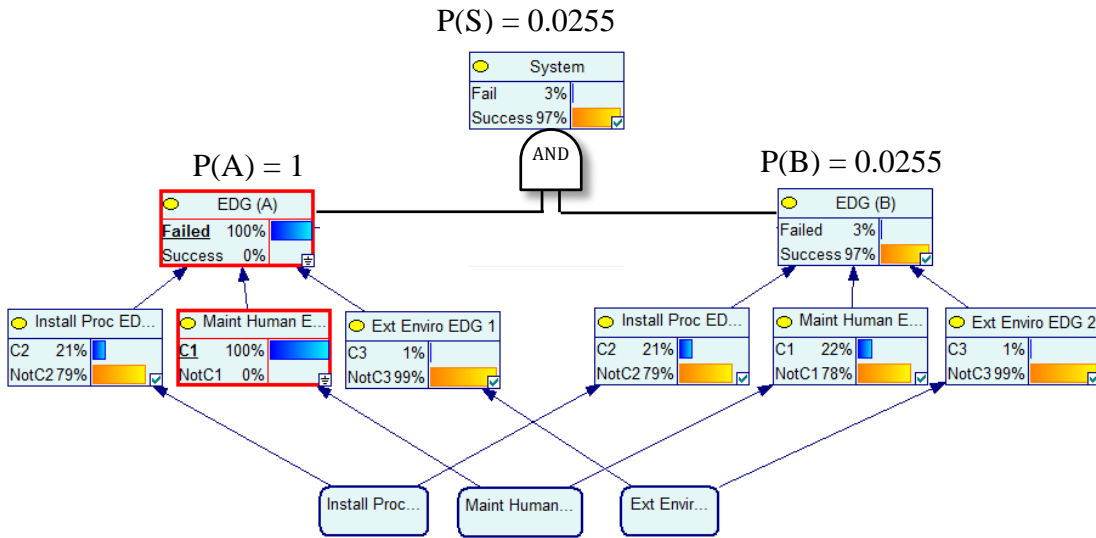
The system failure probability is  $3.810 \times 10^{-2}$  which is higher than the AFM estimate and PAFM of  $6.120 \times 10^{-3}$ . This is because on pump failure, the cause may be due to the maintenance team or external environment which is a cause shared by the EDGs. The Bayesian Network propagated this possibility to the EDGs and increases their probability of failure.

### 7.8.2. Knowledge of Failure Cause

Where the failure cause is known, the system equation can be updated by instantiating

the local cause node as true. Figure 57 shows the event assessment for example 1 where component A failed due to Maintenance Human (MH) error, while causes IP and EE at EDG1 remain unconfirmed. The evidence has once again propagated through the Bayesian Network and updated the probability that a cause condition exists at component B.

Of note, the probability that the cause MH exists at EDG 2 is 0.22 which is an increase from 0.027 due to the addition of the common cause condition from component A, where  $\eta_{MH}^{[E]} = 0.2$ . This is the expected result calculated algebraically in section 7.5.2 and appendix 4.



**Figure 57: Event assessment for component A failing from cause MH using GDM**

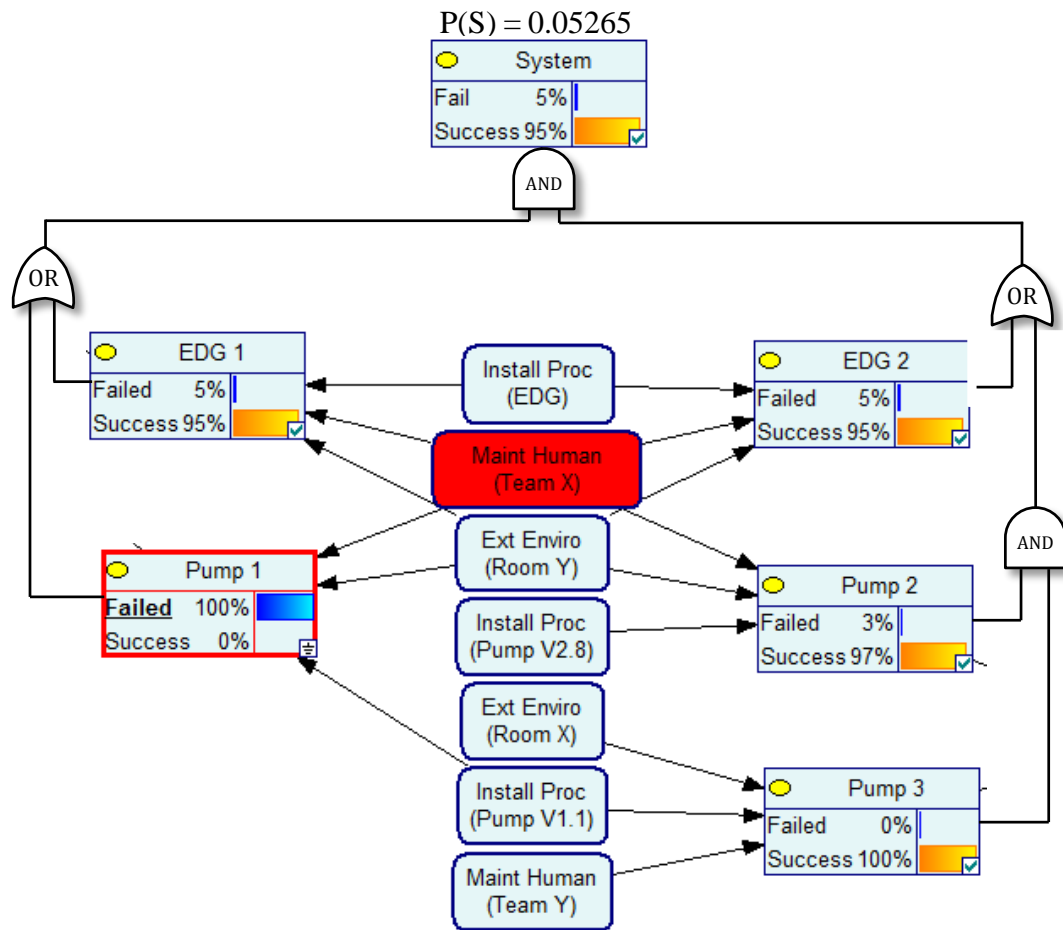
The probability of system failure given knowledge about each cause, is provided in Table 46. The estimates for each cause occur either side of the marginal distribution

when the cause is not known.

**Table 46: Event Assessment for Example 1 with different failure causes**

Cause	$P(S)$	System Failure Probability
Unknown	$P(S A)$	0.0240
Install Procedure Error	$P(S A, C_{IP})$	0.0225
Maintenance Human Error	$P(S A, C_{MH})$	0.0255
External Environment Shock	$P(S A, C_{EE})$	0.0663

For example 2, an event assessment where Pump 1 has failed due to a Maintenance Human cause condition is shown in Figure 58.



**Figure 58: Event assessment for component P1 failing due to EE using GDM**

The EDG1, EDG2, Pump 1, and Pump 2 all share the same maintenance team. With Pump 1 failing due to a human error from that maintenance team, the Bayesian Network now propagates this evidence, and increases our belief that EDG1, EDG2 and Pump 2 could fail.

The event assessment of Pump 1 failing due to installation procedure is shown in Figure 59.

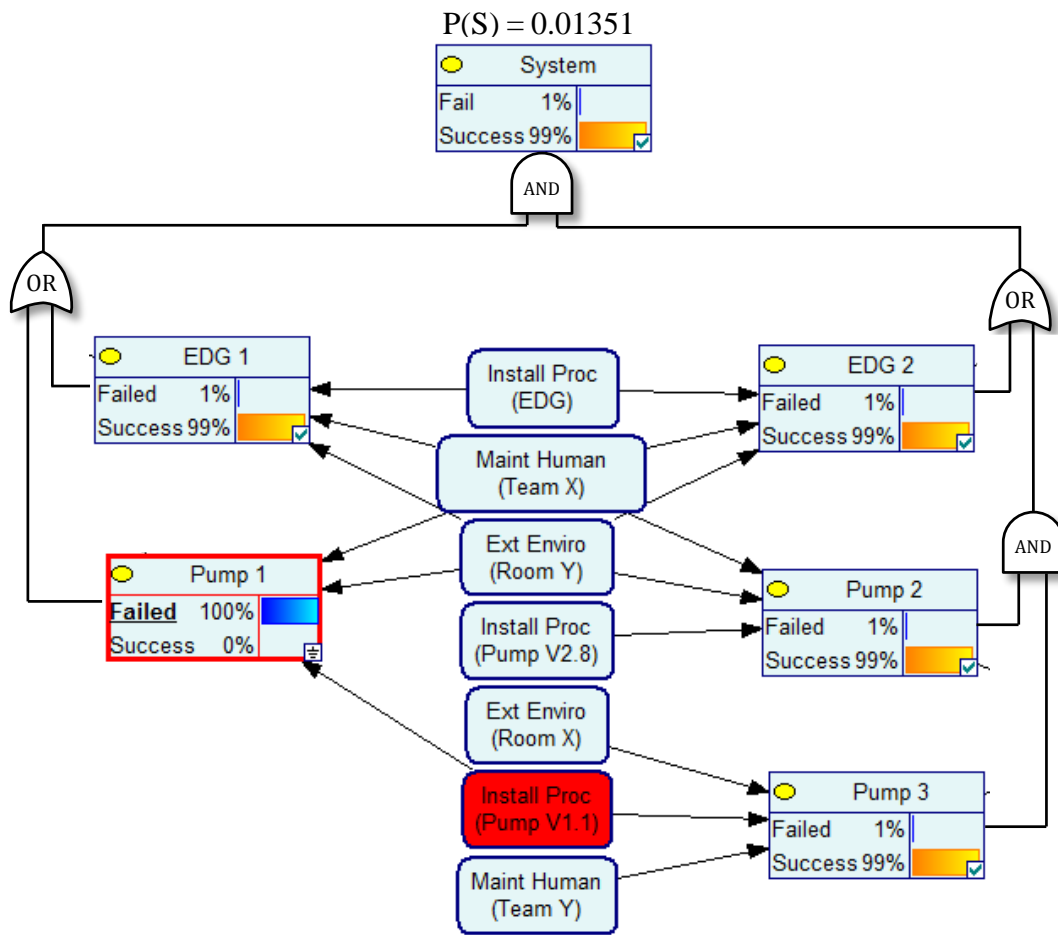


Figure 59: Event assessment for component P1 failing due to IP using GDM

With an error identified in the Pump V1.1 Installation Procedure, it is clear that this has

no effect on the EDGs or Pump 2 which was installed using a different procedure. Therefore the only Pump 3 has an increased probability of failure due to it sharing the same procedure.

The probability of system failure for each possible cause for Pump 1 is contained in Table 47. The table includes a comparison between the PAFM and the GDM estimates. The GDM estimates are higher because they include many more dependencies. The PAFM estimates cannot account for shared dependencies between components (or component failure modes, if that was included in the model).

**Table 47: Event Assessment for Example 2 with different failure causes**

<b>Cause</b>	<b><math>P(S)</math></b>	<b>PAFM</b>	<b>GDM</b>
Unknown	$P(S A)$	6.120e-3	0.03810
Install Procedure Error	$P(S A, C_{IP})$	6.100e-3	0.01351
Maintenance Human Error	$P(S A, C_{MH})$	6.053e-3	0.05265
External Environment Shock	$P(S A, C_{EE})$	6.154e-3	0.03789

### 7.8.3. *Uncertain Knowledge of Failure Cause*

Reports on CCFs failures may be difficult to interpret the characteristics of the failure. The impact methodology was created to allow for the uncertainty in a CCF failure to be captured correctly within a failure database. The Bayesian Network also allows for interpretation uncertainty when conducting event assessments.

When a variable is not absolutely known and the observer has a distribution over the variable's possible values this is often referred to as 'Uncertain Evidence' or 'Soft



Evidence'. There are two types of uncertain evidence, 'Virtual Evidence' and 'Jeffrey's Rule' (Darwiche 2009, p.39). Virtual Evidence is where the uncertainty of the variable has been considered independently of the currently held beliefs, and as such has also been labeled the 'nothing else considered method' (Darwiche 2009, p.39). Jeffrey's Rule is an estimate of the random variable probability after evidence has been applied by placing limitations on the posterior belief of the variable. Jeffrey's Rule has also been labeled the 'all things considered' method. This section will only discuss Virtual Evidence.

Virtual Evidence can be applied where the analyst is unsure of a node's state, but has new evidence to change their belief. Such evidence during event assessment might be available from an initial incident investigation where the failure cause cannot distinguish between a design failure and a manufacturing failure.

The analyst must estimate the odds that each node is true over the other node states. For example, a belief that a cause condition node is 50:50, provides no new evidence to the Bayesian Network and therefore the marginal distributions of the nodes will not change. However, if the analyst believes a node state is 90:10 based only on new evidence, then the Bayesian Network will combine this estimate with the historical evidence which was used to quantify the parameters and provide an updated estimate of the nodes across the network.

For example, the circumstances behind the failure of Pump 1 in example 2 are not clear. The analyst believes there is a 30:70 odds that the failure was due to cause Maintenance Human, and a 70:30 odds that the failure was due to Installation Procedure. The external environment cause has been ruled out.

Figure 60 shows the Bayesian Network where evidence has been applied to the Room Y node so that no external environmental condition exists. With only two options for a failure cause, there is a probability of 0.71 that the failure was caused by Maintenance Human. Figure 61 shows the Bayesian Network including the virtual evidence which captures the analyst's odds about the failure cause. Given this new evidence, the probability that the failure was caused by Maintenance Human has dropped to 0.32.

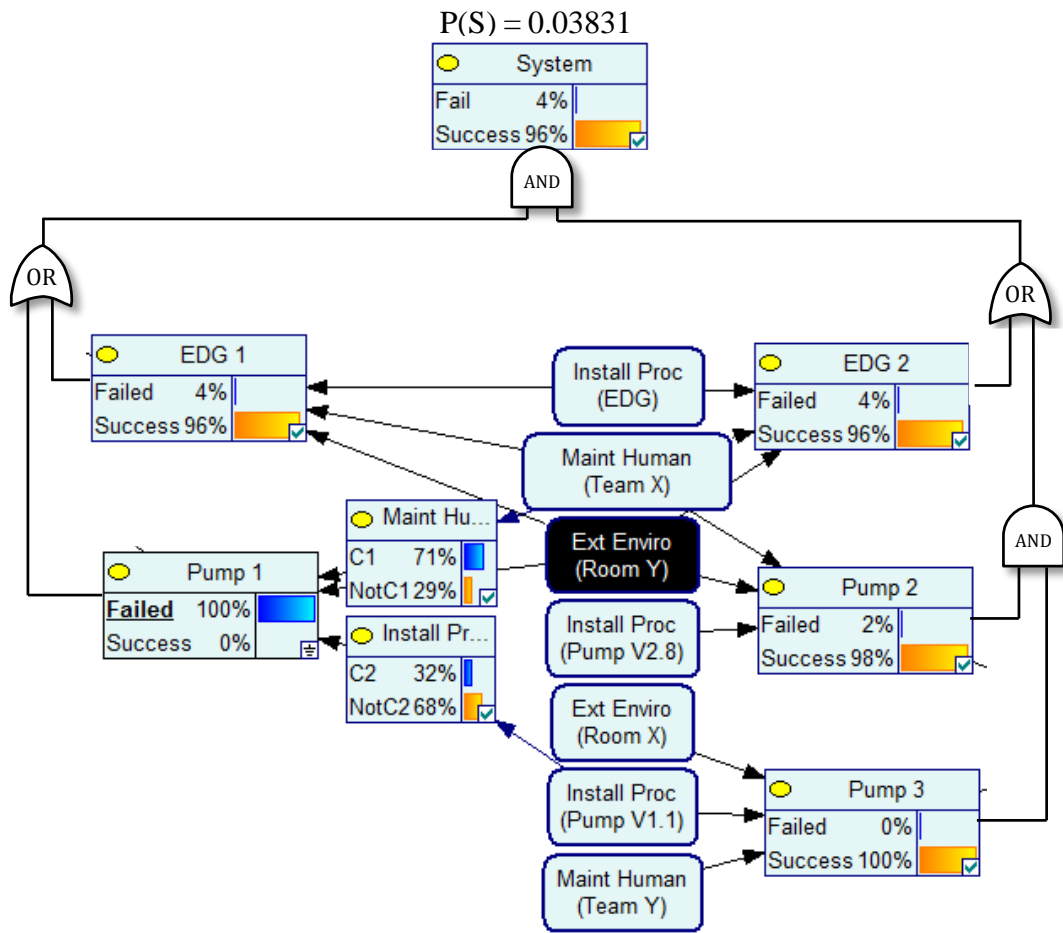


Figure 60: Event assessment for component Pump 1 prior to applying virtual evidence

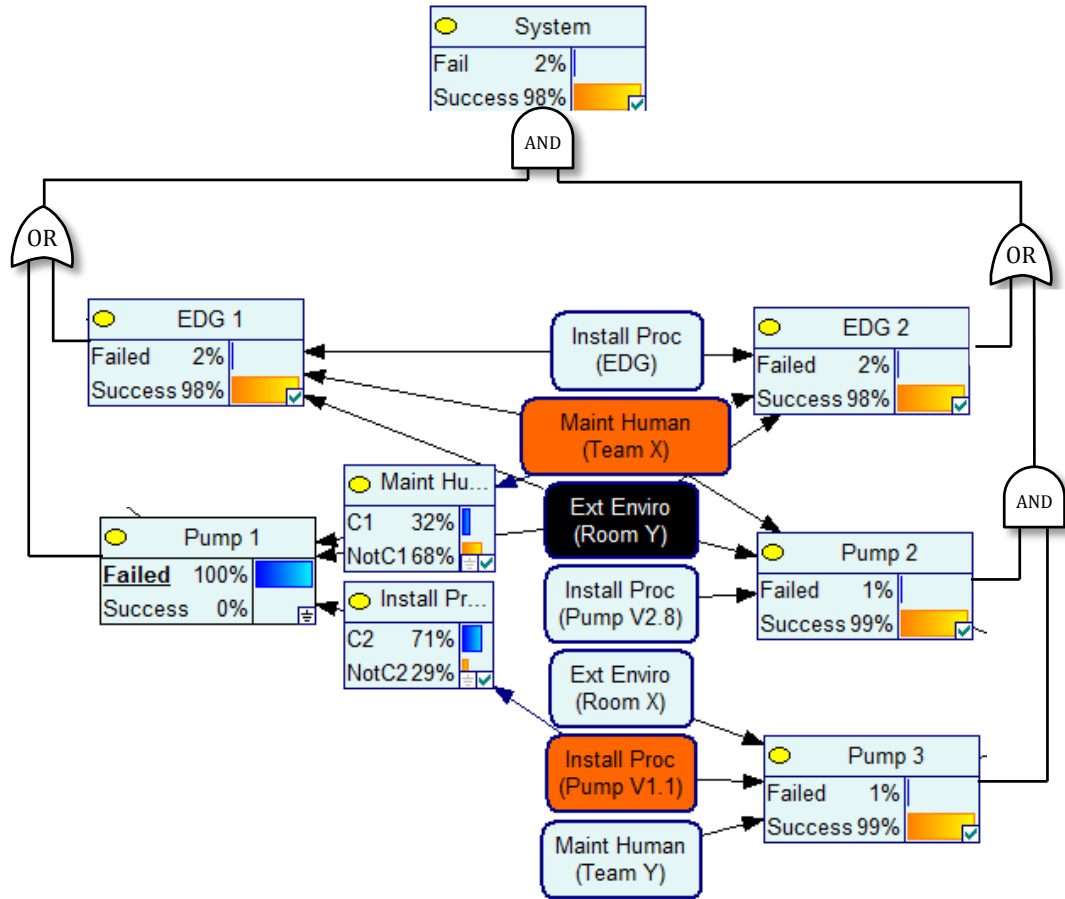


Figure 61: Event assessment for component EDG 1 after applying virtual evidence

### 7.9. Data Collection Requirements

The data collection requirements are the same as for the Partial Alpha Factor model, and discussed in detail in 0. In summary the desired data inputs are:

- The failure cause is recorded for single and multiple failure events.
- The potential coupling factors through which failure propagation could occur is recorded.

- A mutually exclusive, one to one relationship between failure causes and coupling factors exist.
- The size of the CCCG for single failures is recorded.

In addition to those considerations discussed in chapter 6, it is likely that the GDM model will require an assessment of whether failure events are a lethal shock. This data requirement will be discussed as part of extensions to the proposed GDM model. This data is already recorded as part of the NRC failure database.

#### ***7.10. Model Assessment***

The General Dependency Model aims to provide a cause based CCF model which can be quantified using data from general and specific failure databases. Furthermore the GDM aims to extend CCF model past traditional limitations such as asymmetrical components, customized coupling and cause strengths and the ability to conduct cause based event assessments in an efficient way.

This section will briefly describe the advantages, limitations and conduct a comparison of the GDM with other proposed CCF models.

##### ***7.10.1. Model Advantages***

The GDM has a number of benefits compared to other CCF models:

- Allows greater resolution on event assessments.

- Event assessments can be conducted easily without manual quantification of random variables.
- Backward compatible to an equivalent AFM, for small size CCCGs.
- The model parameters have a physical interpretation to assist with comparison and engineering assessments.
- The model can use target system failure rates combined with generic alpha factors to quantify parameters.
- The model allows for a rich description of the target system features such as cause and coupling factor features and defenses.
- The model can account for asymmetrical components.
- The model can account for system specific mitigation defenses through direct adjustment of the parameters.
- The model incorporates failure data in estimation of its parameters.
- The Bayesian Network does not modify the structure of the PRA fault tree and basic events.

### ***7.10.2. GDM Limitations***

Due to the increased complexity of GDM and an attempt to model the physical process of CCF, the GDM has the following limitations:

- The use of a Bayesian Network requires a software tool to account for its complexity.
- The model parameters are not directly observable.
- The model parameters may not be fully specified by the observable data metrics.
- Many of the failure causes will have no observed CCF events and therefore the parameter estimates rely more on the prior knowledge.
- In order to use the NRC failure databases, it must be assumed that each component within the CCCG for the observed failure has the potential for propagation of that cause through a coupling factor. This assumption may not be true and will produce an optimistic estimate.
- Impact vector mapping is still required if data is from a different size CCCG.
- Is not equivalent to AFM when outside the following requirements:
  - By assuming each failure cause event is independent, the model accounts for occasions where failure cause events intersect. This means the resulting failure rate is slightly less than the AFM estimates.

- There are large Common Cause Component Groups.

A limitation of this research has been the ability to test the model against data where the assumptions are clear. Therefore it will be left to future research, with the assistance of an agency such as the NRC, to conduct the following tasks:

- Test the model for larger CCCGs against failure data where the assumption of perfect symmetry can be established.
- Propose a failure taxonomy which minimizes subjectivity of classifications, and establishes a one to one relationship between coupling factors and failure causes.
- Refine the procedures for considering asymmetrical components and their common cause condition probabilities.

### ***7.10.3. Compare Against Model Criteria***

Table 30 provides a comparison of the GDM features compared to previously proposed models.



**Table 48: Assessment of the GDM compared to previous CCF models**

	General Dependency Model	Partial Alpha Factor Model	Basic Parameter	Beta Factor	Partial Beta Factor	Alpha Factor Model	Binomial Failure Rate Model with Lethal Shocks	Common Load	Reliability Cut Off	Influence Diagram	Bayesian Network
Feature Description	GDM	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Explicitly Models System Features	GDM	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Models failure cause	Y	Y	N	N	P	N	N	N	P	Y	Y
Models failure cause defense	P	N	N	N	Y	N	N	N	Y	N	P
Models coupling factor	Y	Y	N	N	P	N	N	N	P	Y	N
Models coupling factor defense	P	P	N	N	Y	N	N	N	Y	N	N
Models deeper causal levels	Y	N	N	N	N	N	N	N	N	N	Y
Models cause condition / shock	Y	N	N	N	N	N	Y	Y	N	N	Y
Models multiplicity of failures within CCCG	Y	Y	Y	N	N	Y	Y	Y	N	N	Y
Models includes consideration for rectification period	N	N	N	N	N	N	N	N	N	N	N
Common Cause Component Grouping Characteristics		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Model non-symmetrical but similar components within the same CCCG	Y	Y	N	N	N	N	N	N	N	N	Y
Model different components within the same CCCG	Y	N	N	N	N	N	N	N	N	N	Y
A component can be part of many CCCGs	Y	Y	N	N	N	N	N	N	N	N	Y
No limit to CCCG size	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Model different failure multiplicities within the CCCG ( $k$	Y	Y	Y	N	N	Y	Y	Y	N	N	Y

Event Assessment Capabilities		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Event Assessment with knowledge of a failed component	Y	Y	Y	N	N	Y	Y	?	N	Y	Y
Event Assessment with knowledge of failure cause	Y	Y	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence - Partial Failures	Y	N	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence- Virtual evidence of cause	Y	N	N	N	N	N	N	N	N	Y	Y
Parameter Estimation		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Impact Vector Method (including method for incorporating	P	Y	Y	P	N	Y	Y	N	N	N	N
Expert estimations (in absence of any data)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account for reliability growth (discount previous failures)	N	N	N	N	N	N	N	N	N	N	N
Update parameters with new evidence	Y	Y	Y	P	N	Y	Y	Y	N	N	N
Incorporate evidence from different sized CCCGs	P	Y	N	P	N	P	Y	Y	N	N	N
Account for CCF which occurred in a different mission time	N	N	N	N	N	N	N	N	N	N	N
Account for CCF data which has artificial separation in time due	N	N	N	N	N	N	N	N	N	N	N
Use system specific failure rate data combined with generic	Y	Y	N	Y	N	Y	N	N	N	N	N
Uncertainty Characteristics for Parameter Estimation		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Does not require distinguish between independent and single	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Failures outside the mission period	Y	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty of shared cause	Y	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty of coupling factor	Y	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty in intervals due to staggered testing	P	P	P	P	N	P	P	N	N	N	N
Partial failures and component degradation	Y	Y	Y	P	N	Y	Y	N	N	N	N
Usability and Cultural Considerations		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Backward compatible to Alpha Factor Model parameters	Y	Y	Y	N	N	Y	N	N	N	N	N
The time investment is no more than the alpha factor model.	P	Y	Y	Y	Y	Y	Y	N	Y	N	N
Automatic parameter estimation is possible from the	P	Y	Y	Y	N	Y	Y	N	N	N	N

## ***7.11. Extensions and Future Development of GDM***

The General Dependency Model discussed within this thesis has focused on proposing a model which could replace the AFM. There are however a number of possible extensions which could be made in order to improve the model's accuracy or flexibility for use in other cases. Furthermore, beyond the research conducted as part of this thesis there are areas which require further work in order to make GDM a mature CCF model.

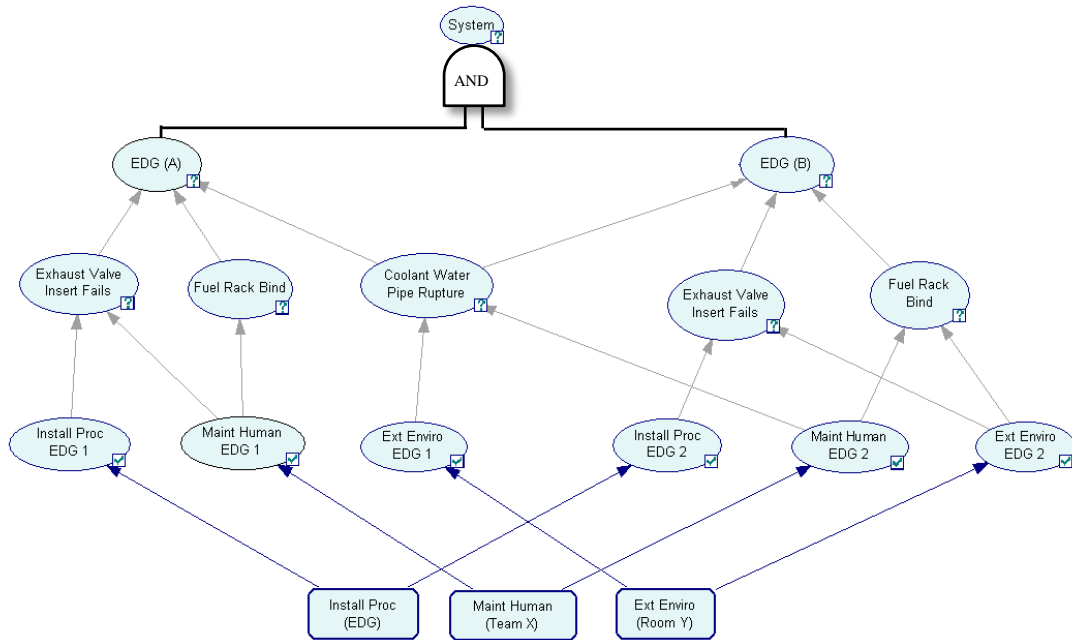
### ***7.11.1. Scalability***

Due to the use of a Bayesian Network, the GDM can model multiple levels of causality. (Kelly et al. 2011) proposed a conceptual CCF model which uses a Bayesian Network to model failure mechanisms such as 'Exhaust Valve Insert Fails' or 'Fuel Racks Bind'.

The General Dependency Model is flexible enough to include such detail by either (1) using the GDM cause condition construct, or (2) using normal Bayesian Network nodes to model probabilistic dependencies. Redefining the GDM cause condition construct is reasonably straight forward, as the definition of cause condition is not confined to a level of causality. However, an alternative approach is to create a network of nodes between the cause condition and the component failure.

Figure 62 shows an example of a GDM model with intermediate nodes, and in this case modeling failure mechanisms similar to (Kelly et al. 2011). This example shows how

the failure mechanism probabilities can be influenced by the presence of a cause condition. Furthermore the failure mechanisms may be local, or they may affect more than one component. For example a coolant water pipe rupture may influence the failure of both EDGs through a cascading failure. A more detailed proposal for physics-based CCF modeling using Bayesian Networks is provided in (Mohaghegh et al. 2011)



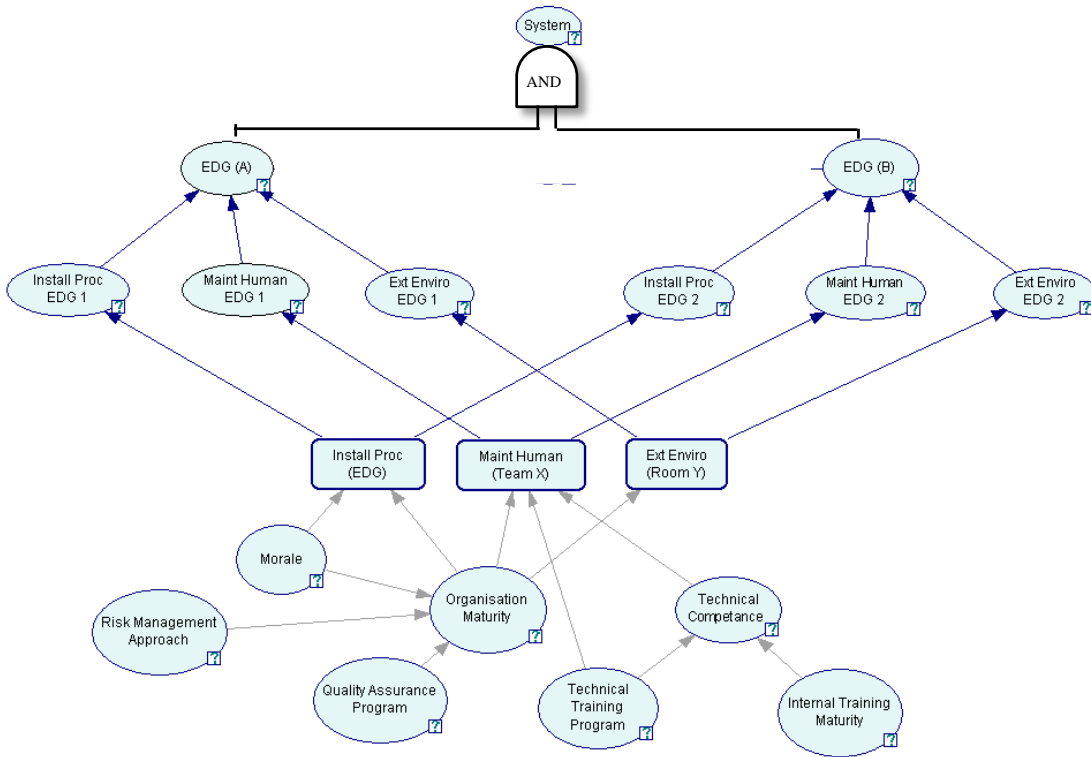
**Figure 62: GDM with intermediate nodes**

In addition to intermediate nodes, it is important to recognize the higher level influences within the system. This may simply model higher levels of coupling as defined in the qualitative assessment. For examples 1 and 2 the location cause condition was defined as being in the same room. However components which are within the same building may also experience CCF due to a shared shock, or the fact that the

components on a particular site are geographically co-located.

In addition to physical traits of the system, the organizational factors may be included within the model. Factors which could be included are the maturity of the organization, quality of an internal training program, common technical training institutes, morale, risk management maturity, common decision makers, etc. For an example of a modeling framework to link human and organizational root causes to a GDM model see (Mohaghegh & Mosleh 2009; Mohaghegh et al. 2009; Mosleh & Goldfeiz 1997)

Figure 63 shows an example where higher level influences have been modeled as parent nodes to the GDM model.



**Figure 63: GDM with parent nodes**

The Bayesian Network construct provides an excellent means to add additional layers to modeled causality and dependencies between basic events in the PRA. This will be limited by the ability to quantify the additional nodes of the network. Where specific research develops such models, it may be added to GDM structure.

### ***7.11.2. Lethal Shocks***

The GDM does not explicitly model the relationship of multiple failure events. Instead it treats these events as a symptom of each component being tested by the existence of a shared cause condition. This is analogous to the Binomial Failure Rate Model.

Using the  $\alpha_{2,i}$  parameter in estimating the GDM parameter guarantees that the GDM model will be faithful to empirical relationships for smaller component groups. For larger component groups the Binomial Failure Rate model was found to be inaccurate, and therefore the concept of lethal shocks was introduced (Mosleh et al. 1998). The Binomial Failure Rate Model and the General Dependency Model are different in the following ways:

- BFRM assumes all components are symmetrical. GDM recognizes asymmetry in both the dependency and the cause strength.
- BFRM assumes that all components receive the same shock. GDM recognizes that the shock may or may not propagate to other components of the coupled components.
- BFRM is calculated without consideration for the failure causes and specific coupling factors. GDM is customized to the cause characteristics
- BFRM splits the failure frequency into independent and dependent. GDM explicitly calculates single failure frequency for only those failure causes which have no coupling factor.

In particular the larger the CCCG, the more likely the assumption of symmetry is violated. The ability for GDM to more accurately model features of the target system may allow for accurate estimations for large similar component groups, however this cannot be validated without the support of industry experts to assist in the isolation of a number of datasets with known assumptions about symmetry.

It may be required that GDM is extended to model lethal shocks separately to non-lethal shocks. The NRC failure database already classifies failure events in terms of lethality. The method of integration into GDM is left for future research but options available include:

- *Multi-state nodes.* Currently the cause condition nodes are ‘Cause’ or ‘No Cause’. A third state may be introduced as ‘Lethal Cause’.
- *A separate network.* A simplistic network could be added to the cause conditioning modeling to represent lethal shocks.

### ***7.11.3. Consistency of Asymmetrical Components***

The GDM system analysis methodology and event assessment procedures have shown how asymmetrical components may be modeled with the GDM. While the model is already capable of this, improvements can be made to the procedures for ensuring consistency, particularly when there are more than two asymmetrical components that require integration. The consideration for such integration will be briefly discussed here, however a detailed analysis is left for future research.

The key to integrating asymmetrical components is the separation of the existence of a cause condition and the reaction the component has to this cause condition. Similar concepts are included within the inference and shock models previously proposed.

Considerations required to formulate a consistent approach for asymmetrical



components are:

- Resolving the issues around the definition of a cause condition such as those discussed in section 7.2.4 which include:
  - Interpretation of multiple shocks during a mission period;
  - Interpretation of shocks with different intensities;
  - Accounting for non-impulse shocks (cause conditions) which remain for many mission periods.
  - Accounting for simultaneous multiple cause conditions.
- The possibility of solving for the third GDM parameter using information from two asymmetrical components.
- The requirement to minimize error when estimating GDM parameters using data from more than two asymmetrical components.

#### ***7.11.4. Uncertain Evidence In Event Assessment***

The use of uncertainty in event assessment was described in section 7.8.3 with the caveat that the virtual evidence estimate must be independent of evidence already used to quantify model. This estimate might be difficult to make as the expert is required to put aside all previously held beliefs about the frequency of such events and objectively quantify their estimate based on the new evidence.

The alternative measure is to use Jeffrey's Rule which is an estimate of the posterior values, and the remainder of the joint probability distribution is updated such that it meets this constraint. This estimate is also problematic for the experts because they are combining numerous sources subjectively including informative priors, evidence from other parts of the model, historic evidence used to quantify parameters and now the new evidence from the event of interest. Virtual evidence is called the 'nothing else considered' method. Jeffrey's rule is called the 'all things considered' method (Darwiche 2009).

There exists a need for an 'audit of things considered' method. Experts can estimate the quantity in a metric they are comfortable with, and be explicit in the evidence they have considered in their estimate. Any evidence overlap between the expert's estimate and what has already been considered in the model can now be treated. For example an estimate may include information about the historic causes of a failure, and the evidence found from the failure investigation but not consider evidence which has been entered about another component. This auditing technique could list all sources of evidence which have been used to formulate the posterior distribution of the unit of interest, and the expert can tick whether they have knowledge of that evidence and if it was considered in their estimate. Any information which is not within the audit can be entered as new evidence.

This topic will be recommended for future research.

#### ***7.11.5. Failure Taxonomy Development***

As detailed in chapter 4, the current CCF taxonomy used within the NRC failure databases has ambiguity when making inference about which failure causes could propagate through which coupling factors.

The development of an unambiguous failure classification taxonomy must be conducted in consultation with the industry for which the CCF model will be used.

#### ***7.11.6. Prior Estimations***

A convenient and defensible method of creating a prior distribution is to use a population variability function, as detailed in section 6.5.3. This prior can overcome problems associated with parameter estimates which equal zero and can provide significant bounds on the possible values which a parameter may take.

The development of such priors is dependent on the failure taxonomy and industry. The development of such priors is left as a future task.

#### ***7.11.7. Unbiased Estimators***

The estimators for GDM may be biased for the following reasons:

- Data is collected as if causes are mutually exclusive, however GDM assumes that causes are independent. Therefore component failure rates are less than

those observed directly from data. This problem becomes prominent only when probability of failure is greater than 0.01.

- Data collected is subject to different testing schemes. This particularly affects components in standby where their functional state may be tested through a staggered or non-staggered theme. Furthermore once a failure is found the maintainer may test other components not part of the schedule. Different testing schemes provide different bias within the data. The GDM estimators need to be developed for each testing scheme.

#### ***7.11.8. Integration of Existing Parametric Models***

Section 7.6.4 described the suitability of existing models to quantify GDM parameters. Specifically, human reliability assessment models exist which may enable the estimation of GDM parameters for human related causes. A literature review and methods for integrating such models are left as a future task.

#### ***7.11.9. Software and Procedure Development***

GDM has a reasonably simple analysis methodology, but requires software to calculate the required outputs. The system analysis procedure essentially involves a qualitative assessment for coupling factors, the classification of each basic event into those coupling factor categories, and the quantification of three parameters for each

dependency.

The system analysis procedure does not require an understanding of Bayesian Networks to complete, and may be completely conducted through a graphical user interface which only displays the event trees and fault trees of the original PRA.

In event assessment there are significant benefits to displaying the updated probabilities of the cause conditions. This need not be displayed through a Bayesian Network, but may be presented to the user as an event assessment dashboard.

Furthermore, in conducting tasks such as the qualitative assessment, event assessment or use of uncertain data, checklists and procedures may be developed such that the task is intuitive and easy to conduct.

Previous CCF models have been proposed which overcome many of the disadvantages of the popular AFM. However they have failed either through difficulty in obtaining/incorporating data or through the complexity of the analysis process. The analysis of CCF using GDM requires software. It is left as a future task to develop such software that the complexities of constructing the Bayesian Network are shielded from the user and only a simple, intuitive, software guided analysis process remains.

## Chapter 8: Conclusion

### 8.1. Introduction

The aim of this final chapter is to discuss the topics covered within this thesis, provide a summary of the proposals made, review the advantages and limitations of each proposal and discuss the extent to which the research objectives have been met. Finally areas for further research into common cause failure will be identified.

### 8.2. Review of research objectives and goals

The goal of this research was to develop a comprehensive CCF analysis methodology that enhances the discipline of PRA to conduct the following tasks:

- CCF analysis in the PRA of operating system in particular allowing the analysis to recognize system specific features such as inter-component dependencies and CCF defenses.
- CCF analysis of systems in design, in particular allowing the effect of different design decisions to be quantified in the absence of plant specific data.
- CCF analysis in Event Assessments, in particular the ability to include characteristics of the failure event which allow an assessment of the probability for CCF.

To support this goal, research objectives were specified and achieved as follows:

*Propose a unified understanding of the definition and scope of CCFs.* This objective has been met through a literature review covering the ambiguity in CCF definition, an analysis of the attributes of CCF and through a proposed redefinition of CCF which addresses the identified issues and attributes.

*Propose a failure data taxonomy consistent with the unified CCF definition which can enable cause based CCF models.* This research objective has been met through the analysis of the CCFDB for suitability to support the cause based CCF models. It was found to be insufficient and an alternative failure classification scheme was developed.

*Propose a cause based, data informed CCF model.* This research objective has been met through the proposal of two new CCF models. The first is based on an extension of the Alpha Factor Model with minimal changes to the analysis process. The second model is based on a Bayesian Network which requires modification to traditional analysis process.

*Propose a comprehensive and scalable analysis process.* This research objective has been met by analyzing the current CCF analysis process for both system design and event assessment. For each proposed CCF model, a new analysis process was developed and demonstrated.

### 8.3. Common Cause Failure Definition

The definition of Common Cause Failure has many different meanings depending on the context, industry and whether the term is being used to describe modeling or the general description of a failure phenomenon. Consequently there is a lack of consensus regarding a definition which is consistent in a modeling and phenomenon context.

Using the CCF definition from NUREG/CR-5485 as a baseline, the following areas were found to require further clarification.

#### **CCF defined within a PRA modeling concept.**

Previous CCF definitions have described CCFs as a phenomenon without reference to modeling. However CCF modeling is only conducted on dependencies not explicit within the PRA model. As models become more detailed in their treatment of dependencies, the definition of a CCF for that model also changes. This means that two models for the same scenario may have different meanings of the term CCF.

A definition is required to understand why multiple components of a system may fail due to unforeseen circumstances, and to support data collection activities and define the scope of CCF modeling. Therefore the decision as to whether to have the definition of CCF dependent on the PRA model gives two options:



- *Exclude reference to modeling within the definition.* This is the status quo, which provides a broad definition of CCF which allows for explicit or implicit modeling. This definition would mean any dependent failure would be considered a CCF regardless of whether it is explicitly modeled or not. CCF modeling would only model a fraction of CCFs in a PRA model, and the CCFDB only collects a portion of CCF events due to it ignoring more explicit dependent failures.
- *Include reference to modeling within the CCF definition.* Including modeling as part of the CCF definition implies that CCF is a modeling concept used to represent a sub-set of dependent failures. This means an observed event cannot be classified as a CCF unless there is knowledge of the PRA model used to represent the failure.

It is this problem which creates the most confusion about the definition of CCF. Without resolving this issue there is no definitive way to differentiate a CCF compared to a commonly understood dependent failure which does not help support data collection activities or the progression of CCF modeling. This thesis proposes that the definition for CCF includes reference to whether the failure dependency is explicitly or implicitly modeled.

### **Simultaneity**

The undesirable effect of CCF is when two or more components are unable to perform

their function at the same time. All CCF modeling has focused on the probability of component failure during a defined mission period. An extension to the concept of simultaneity is consideration that the probability of two or more components being unavailable to perform their function at the same time is linked to how quickly a component is repaired. A small addition to the definition has been proposed to clarify this.

### **Proposed Definition**

The following amended CCF definition is proposed.

A CCF event consists of component failures that meet five criteria:

- (1) two or more individual components fail or are degraded, including failures during demand, in-service testing, or deficiencies that would have resulted in a failure if a demand signal had been received;
- (2) component failures occur within a specified period of time such that multiple components are unable to perform their intended function or success of the PRA mission would be uncertain;
- (3) component failures result from a single shared cause and coupling mechanism;
- (4) a component failure occurs within the established component boundary; and
- (5) the dependency between components has not already been explicitly modeled.

#### **8.4. CCF Failure Taxonomy**

To enable the quantification of the proposed cause based CCF models, failure events must be classified such that there is a clear relationship between the failure cause and the coupling factor. The Partial Alpha Factor needs to know the coupling factors a failure could have propagated through. The General Dependency Model quantifies the strength of each coupling factor using data from failure causes. This aim is best achieved if there is a mutually exclusive, one to one relationship between the failure cause and coupling factor.

An assessment was conducted of the CCFDB and it was found that there is insufficient correlation between the failure cause and coupling factor categories for it to be used in its current form.

The features of a suitable taxonomy were described and a failure event taxonomy was proposed based on the existing coupling factor categories, for review by industry experts.

#### **8.5. CCF Model – Partial Alpha Factor Model**

##### **8.5.1. *Overview***

The Alpha Factor Model is the most popular quantitative model in use, and is well

supported by the data collection and classification systems available. Therefore the Partial Alpha Factor Model aims to provide a cause based model which uses the same analysis methodology and data sources as the AFM.

The PAFM is quantified using two parameters, the partial alpha factors ( $\alpha_{k,i}$ ) and the gamma factors ( $\gamma_i$ ). The partial alpha factors represent the multiplicity of failures in a common cause component group, given a failure has occurred from cause  $i$ . The gamma factors represent the portion of failure events which are due to cause  $i$ . Given an assessment of the coupling factors between two components, the ‘assessed alpha factors’ may be calculated as:

$$\alpha'_k = \sum_{i=r} \gamma_i \alpha_{k,i}$$

- $r$  = *the coupling factors shared by the components within the CCCG being analysed,  $r \subseteq \{1,2,3,\dots,w\}$ .*
- $\alpha'_k$  = *the assessed alpha factor. This is the system alpha factor which only considers the coupling factors shared by the components within the CCCG where  $2 \leq k \leq m$*

Where the components share all coupling factors, the assessed alpha factor is equivalent to the AFM parameters.

The analysis process is the same as the AFM with the following significant differences:

- A Common Cause Component Group (CCCG) is formed around symmetry of shared causes. This means a component may be a member of multiple CCCGs.

- Common Cause Basic Events are created for each CCCG, and result in more than would have been required using the AFM.
- Parameter estimation occurs for each cause, instead of at the failure mode level.

### 8.5.2. *Advantages*

The primary advantages of the PAFM is the ability to customize the CCF model to relevant coupling factors between components and conduct event assessments with knowledge of the cause whilst using a familiar AFM methodology.

Specifically the advantages of the PAFM over other CCF models are summarized below:

- Allows greater resolution on event assessments.
- Backward compatible to an equivalent AFM.
- Intuitive extension to the AFM analysis methodology.
- A ratio model allowing the use of target system failure rates.
- Can reward target system defenses that decouple dependencies.
- Can use AFM for system analysis, and the PAFM used for event assessment.
- Using the assumptions contained in section 6.8.1, the PAFM can be calculated from the CCFDB.
- PAFM parameter estimates will be no worse than if the PRA used the AFM.

A detailed comparison of the PAFM features against the other CCF models is provided

in Table 30 on page 175.

### **8.5.3. *Limitations***

Due to the PAFM using an AFM methodology and the nature of CCF data, the PAFM has a number of limitations:

- The description of the target system features such as cause and coupling factor features and defenses is limited.
- Many of the failure causes will have no observed CCF events and therefore the parameter estimates rely more on the prior knowledge.
- As per the AFM, it is difficult to model components with different failure probabilities within the same CCCG (symmetrical failure probabilities)
- In order to use the CCFDB, it must be assumed that each component within the CCCG for the observed failure has the potential for propagation of that cause through a coupling factor. This assumption may not be true and will produce an optimistic estimate.
- Impact vector mapping is still required if data is from a different size CCCG.

## 8.6. CCF Model - General Dependency Model

### 8.6.1. Overview

The General Dependency Model aims to provide a feature rich, cause based CCF model that can be quantified using a CCF database. Significant objectives include:

- The ability to conduct event assessments using uncertain evidence and knowledge of the failure cause;
- The ability to model features of the target system including cause frequency, coupling factor strength and system defenses;
- The ability to model asymmetrical components; and
- The ability to quantify using impact vectors.

The model uses three parameters for each cause, the component fragility ( $p_i$ ), the cause condition probability ( $Q_{E,i}$ ), and the coupling factor strength ( $\eta_i$ ). These parameters allow the separation modeling of cause conditions (i.e a flood), to the reaction equipment have to that failure cause (i.e probability of failure given a flood). The coupling factor strength describes the ability of system to propagate the cause condition, not the failure (i.e a flood propagating to a different buildings). This allows for the modeling of asymmetrical components and gives greater flexibility in modeling features of the target system.

The structure of GDM uses a Bayesian Network. This allows for a flexible model which can easily propagate evidence through the PRA model and give flexibility in integrating additional factors such as higher levels of dependency or organizational factors.

The implementation of GDM requires a soft solution. Using such a solution allows for a CCF analysis process which focuses on the analyst describing the features of each component, instead of changing the PRA structure. Bayesian Network software also allows for a very straightforward event assessment procedure and a rich graphical interface to interpret event assessments results.

### **8.6.2. Advantages**

The advantages of the GFM over other CCF models are that it:

- Allows greater resolution on event assessments.
- Event assessments can be conducted easily without manual quantification of random variables.
- Backward compatible to an equivalent AFM, for small size CCCGs.
- The model parameters have a physical interpretation to assist with comparison and engineering assessments.
- The model can use target system failure rates combined with generic alpha factors to quantify parameters.
- The model allows for a rich description of the target system features such as cause and coupling factor features and defenses.



- The model can account for asymmetrical components.
- The model can account for system specific mitigation defenses through direct adjustment of the parameters.
- The model incorporates failure data in estimation of its parameters.
- The Bayesian Network does not modify the structure of the PRA fault tree and basic events.

A detailed comparison of the GFM features against the other CCF models is provided in Table 48 on page 271

### ***8.6.3. Limitations***

Due to the increased complexity of GDM and an attempt to model the physical process of CCF, the GDM has the following limitations:

- The use of a Bayesian Network requires a software to account for its complexity.
- The model parameters are not directly observable.
- The model parameters may not be fully specified by the observable data metrics.
- Many of the failure causes will have no observed CCF events and therefore the parameter estimates rely more on the prior knowledge.

- In order to use the NRC failure databases, it must be assumed that each component within the CCCG for the observed failure has the potential for propagation of that cause through a coupling factor. This assumption may not be true and will produce an optimistic estimate.
- Impact vector mapping is still required if data is from a different size CCCG.

#### ***8.6.4. Future Work for GDM***

The General Dependency Model requires further development in order to become a usable capability by PRA practitioners. Furthermore, a number of extensions are recommended for investigation to enhance the model's capabilities.

A limitation of this research has been the ability to test the model against data where the assumptions are clear. Therefore it will be left to future research, with the assistance of an agency such as the NRC, to conduct the following tasks:

- Test the model for larger CCCGs against failure data where the assumption of perfect symmetry can be established.
- Propose a failure taxonomy which minimizes subjectivity of classifications, and establishes a one to one relationship between coupling factors and failure causes.
- Produce non-bias estimators considering the assumption of independent events and testing scheme.

- Refine the procedures for considering asymmetrical components and their common cause condition probabilities.

A number of implementation activities and enhancements which require further work are described in detail in section 7.11 and listed below:

- Create a methodology for integrating higher level influences into GDM such as organizational factors.
- Create a methodology for integrating intermediate nodes between cause conditions and the failure of components.
- Investigate the inclusion of lethal shocks and compare against industry data for larger common cause component groups.
- Develop a consistent methodology for the inclusion of asymmetrical components into the GDM. This activity includes:
  - Resolving the issues around the definition of a cause condition such as multiple shocks during a mission and shocks which last multiple mission periods (see section 7.2.4)
  - The possibility of solving for the third GDM parameter using information from two asymmetrical components.
  - The requirement to minimize error when estimating GDM parameters using data from more than two asymmetrical components.

- The difference between a common cause condition on similar items, versus a common cause condition on asymmetrical components.
- Develop a methodology to enable experts to provide evidence through a ‘evidence audit procedure’ detailed in section 7.11.
- Develop a failure cause classification which meets the requirements detailed in chapter 4, and reclassify events in the failure database.
- Using the data within the failure database, develop population variability distributions which can be used as prior distributions for parameter estimation.
- Investigate parametric models which may be used to estimate parameters of the GDM such as human reliability assessment models.
- Investigate a parametric means of adjusting GDM parameters for different levels of defenses.
- Develop software and procedures to automate the system analysis and event assessment tasks such as creation of the GDM structure and population of node conditional probability tables.

### ***8.7. Future Research***

During the conduct of this research, a number of issues with CCF modeling were identified which are recommended for future research. The following activities relate to all CCF models, not just to those proposed within this thesis.

### ***8.7.1. Time Relationship of CCF (Repair Time, Mission Time, Aging)***

Common Cause Failures are dependent on the time between failure in a number of ways. Despite this there is very little consideration for explicit modeling of these relationships which can severely limit CCF modeling's scope and the ability for CCF databases to be used in other industries.

#### **Repair Time**

The undesirable effect cause by CCF is when multiple components fail to provide their function at the same time. Depending on the repair time this could be minutes, hours, even years and independent of the mission period. However all formulations to date have ignored the repair time as a consideration. A proposed topic for further research is to investigate the integration of repair times into the model calculations.

This topic is discussed in more detail in section 3.4.2.

#### **Changes to Mission Time**

In most cases a CCF has been defined as the failure of multiple components during a defined PRA mission time. When this definition is used to collect data it means that a change in mission time definition will change which failures from a database will be considered CCF.

For example if the nuclear industry created a database of CCF using a definition of a

24 hours mission period, then failures which occurred 2 months apart will be classified as single failures. However if NASA wishes to use the database to quantify its CCF model, and defines its mission period of 5 years, then if two failures occurred 2 months apart it would be considered a CCF. The alpha factors for CCF will be much higher for the NASA mission, however there is no way to adjust the nuclear industry database to account for this new mission time.

It is recommended that a method for recording failures is investigated such that the analyst can define a mission period, and recalculate their CCF parameters. This topic is discussed in more detail in section 2.13.

### **Aging**

When studying observed CCF events, many show that after the initiation of a cause condition, the probability of a simultaneous failure depends on the aging characteristic of the condition. For example, if a maintenance error occurs when installation a filter. The component may work perfectly for the first week, but the error has accelerated the time to failure. So the probability of CCF increases after the initiation event.

It is recommended that this aging characteristic be further investigated to determine if it can be used to better estimate CCF rates.

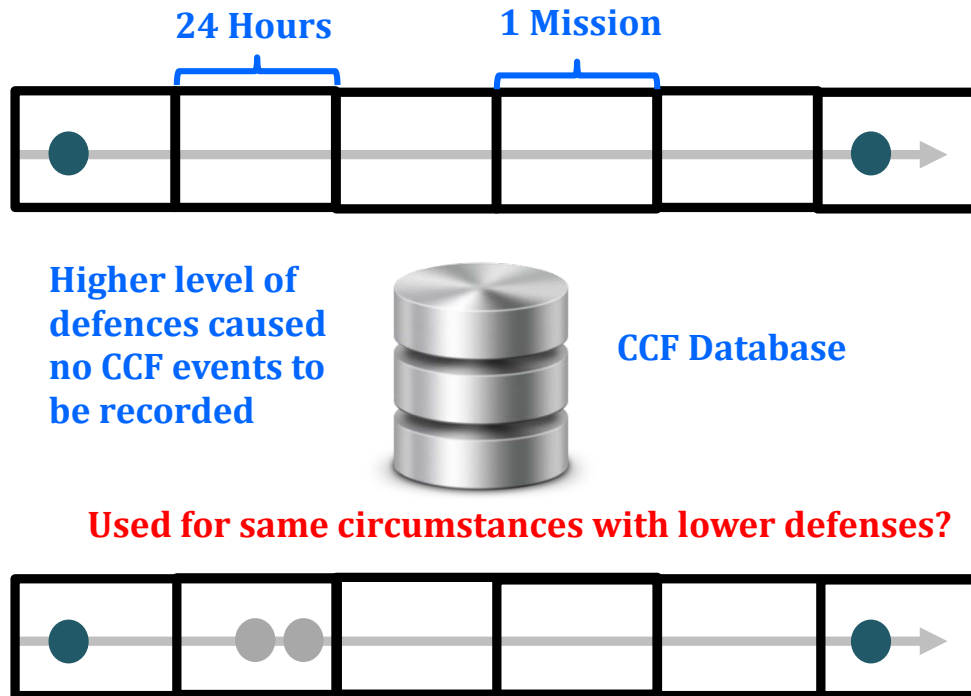
### ***8.7.2. Existing Parametric Model Estimate***

Section 7.6.4 described the suitability of existing models to quantify GDM parameters. Specifically, human reliability assessment models exist which may enable the estimation of GDM parameters for human related causes. A literature review and methods for integrating such models are left as a future task.

### ***8.7.3. Quantitative Defense Modeling***

An ultimate objective of CCF modeling is to quantify the effect that defenses have on cause frequencies, coupling factor strengths and component fragility. This is done qualitatively using UPM. Despite some methods being proposed (Hakansson 2011; Zitrou 2006a), there currently exists no data-informed quantitative model.

In order to provide a data informed model for defenses, consideration must be given to how defenses can be classified when recording failure events. For example, Figure 64 shows how failure event information must include the defense levels to allow inference of target system failure rates and CCF parameters.



**Figure 64: Recording of Failures into CCFDB with consideration for defenses**

It is recommended that further research is conducted to propose a quantitative method for estimating the effect of system defenses, and research into the data requirements to support the model.

### **8.8. Conclusion**

This research has proposed a methodology for modeling CCF with explicit consideration for failure causes and coupling factors. Central to this ability is the proposal of two CCF models. The Partial Alpha Factor Model is an extension of the popular Alpha Factor Model which minimizes additional complexity. The General Dependency Model provides a feature rich CCF analysis capability through an



increased complexity.

In order to quantify cause based CCF models, this thesis reviewed the suitability of the current CCFDB failure event classification taxonomy and found it to be insufficient. An alternative classification method has been proposed which allows inference of the coupling factors to be made with knowledge of the failure cause.

Finally this research investigated the definition of Common Cause Failures, and why consensus on its interpretation has not been achieved. An alternative definition has been proposed.

## Appendices 1: Literature Review of CCF Models

### *1.1. Introduction*

This literature review proposes the following broad classifications of CCF models:

- **Direct Estimate.** This classification includes all models which estimate CCF parameters without assumptions outside directly observed system failure rates.
- **Ratio Models.** Ratio models estimate the percentage of a components failure rate which is caused by common cause. This allows ratios to be estimated from generic data and applied to systems where only a failure rate is known.
- **Shock Models.** Shock models are based on the hypothesis that each component within the CCG undergoes shocks according to a Poisson process. For each component within the CCG the shock is a Bernoulli trial which will fail the component with probability  $\rho$ . This includes models where a probability distribution is assigned to the parameter  $\rho$ .
- **Interference Models.** Interference models also attempt to model the physical phenomena of CCF but without the shock model's assumption of independence. Instead these models predict the number of failures by assuming random variables for component strength and load. When the load exceeds the strength a component is expected to fail. The more intense the load or the more depleted the strength than the higher the probability of failure.

- **Other Models.** This contains all the models which did not fit neatly into the other categories. This included implicit models and a square root bounding method.

This chapter will conclude with a summary of the model classifications and a table comparing the assumptions and limitations of each model.

## ***1.2. Direct Estimates***

Direct estimate models consist of:

- Direct Assessment (Qualitative)
- Basic Parameter Model

### ***1.2.1. Direct Assessment (Qualitative)***

The direct assessment model can be considered as a procedure rather than a modeling technique. (Hirschberg 1985) discussed this approach in detail. It involves using the actual number of demands and the number of observed failures with multiplicity,  $i$ , and estimating the quantities of interest directly from the data set. For the purposes of this model  $i$  is defined as any positive integer, and the quantities of interest for this case is defined as the common cause failure rate or common cause failure probability.

This approach is typically simple and is less dependent upon sound knowledge of any

mathematical or statistical skills. (Anude 1994)

Advantages include:

- The method is simple as there is minimal data required and minimal mathematical knowledge required to determine broad estimates.
- The model is predicated on experience.

Limitations include:

- The model cannot estimate common cause failures for K out of N events for which it does not have data.
- Component symmetry is assumed (ie.  $X_{AB} = X_{BC}$ )
- Does not allow for partial failure or component degradation
- No inference can be made given knowledge of the failure cause.
- The model cannot account for unique system architecture which may contribute or defend against CCF.

### ***1.2.2. Basic Parameter Model***

The basic parameter model was proposed by Fleming, et al. in 1983 (Fleming et al. 1983) and calculates the CCF basic event directly from the data. This estimation is given by: (Mosleh et al. 1998)(Mosleh 1991)

$$Q_k^m = \frac{n_k}{N_k}$$

- $Q_k^{(m)}$  = basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).
- $n_k$  = the number of failure events which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).
- $N_k$  = the number of demands on any  $k$  component in the common cause group.

If it is assumed that each time the system is operated, all of the  $m$  components in the group are demanded, then.<sup>10</sup>:

$$N_k = \binom{m}{k} N_D$$

$$Q_k^m = \frac{n_k}{\binom{m}{k} N_D} \quad \text{non-staggered testing}$$

$$Q_k^m = \frac{n_k}{m \binom{m}{k} N_D} \quad \text{staggered testing}$$

$N_D$  = the number of demands on the system (or time  $T$ )

The total component failure rate can be calculated as the sum of the CCBEs:

$$Q_t = \sum_{k=1}^m \binom{m-1}{k-1} Q_k^m$$

Replacing  $Q_k^{(m)}$  with its estimator gives the following estimator for the total failure rate,  $Q_t$ :

---

<sup>10</sup> This estimator can change depending on the scheme used to test components. This estimator is for non-staggered testing. Other estimators and discussion on testing schemes is provided in NUREG /CR-5485 (Mosleh et al. 1998)

$$Q_t = \frac{1}{mN_D} \sum_{k=1}^m kn_k \quad \text{non-staggered testing}$$

$$Q_t = \frac{1}{m^2N_D} \sum_{k=1}^m kn_k \quad \text{staggered testing}$$

Advantages include:

- The method is simple as there are no intermediate steps in quantifying basic common cause events.
- The model is intuitive.
- Suitable for any amount of redundancy (for which data is available).
- No need to differentiate between independent and common cause failures.

Limitations include:

- The model cannot estimate common cause failures for redundancy configurations for which data is unavailable.
- Component symmetry is assumed (ie.  $X_{AB} = X_{BC}$ )
- Does not allow for partial failure or component degradation.

### 1.3. Ratio Models

Ratio models are based on the hypothesis that system specific estimates for CCF can be made by combining generic average ratio parameters with system specific single/total failure rates (Vaurio 2008). This provides the advantage that ratio models can be estimated from specific data collection activities such as the Common Cause Failure Data Base and applied to areas where CCF data may not exist.

Ratio models have the following advantages:

- There is a direct and intuitive quantity to the model parameters.
- Generic ratio parameters can be calculated from generic data and then applied to plant specific single failure rates. This reduces the data requirements compared to direct estimate models.
- Success data is not required to estimate the model parameters.

The limitations of all ratio models discussed here include:

- The model assumes a transferable empirical ratio between failure rates and Common Cause Failure rate.
- The ratio models described here assume component symmetry.
- No inference can be made given knowledge of the failure cause.
- The model cannot account for unique system architecture which may contribute or defend against CCF.
- Confusion in the interpretation of single failures being modeled as independent failures, particularly when applying impact mapping rules.

Ratio models can be further classified into:

- Component failure ratio models. The ratio is calculated as a count of the number of component failures which occur in different configurations which includes:

- Event failure ratio models. The ratio is calculated as a ratio of different types of events. A single event may have multiple component failures.

### 1.3.1. Beta Factor Model

The Beta Factor Model, proposed by Fleming in 1975 (Fleming 1975), is a component failure ratio model which is one of the most popular where generic data used to estimate parameters are limited. It is still the most commonly used CCF model outside the nuclear industry (Hokstad & Rausand 2008).

The basic parameters can be calculated as (Mosleh et al. 1998):

$$Q_k^{(m)} = \begin{cases} (1 - \beta)Q_t & k = 1 \\ 0 & m > k > 1 \\ \beta Q_t & k = m \end{cases}$$

$Q_t$  = The total failure probability of one component

$m$  = Common Cause Component Group size

$k$  = Number of failed components due to common cause failure

The MLE parameter estimate is (Mosleh et al. 1998)<sup>11</sup>:

$$\hat{\beta} = \frac{\sum_{k=2}^m k n_k}{\sum_{k=1}^m k n_k}$$

$n_k$  = number of events involving  $k$  components in a failed state

$m$  = the number of components within the CCCG

---

<sup>11</sup> This estimator can change depending on the scheme used to test components. This estimator is for non-staggered testing. Other estimators and discussion on testing schemes is provided in NUREG /CR-5485 (Mosleh et al. 1998)



$\beta$  (usually between 0.01 and 0.3) (Anude 1994) is defined as the point estimate of the conditional probability that a unit failure is a Common Cause type. The Beta Factor model uses one parameter in addition to the total component failure probability to calculate the Common Cause failure probabilities regardless of the size of the Common Cause Component Group.

The Beta Factor model is a special cause of the Multiple Greek Letter (MGL) model where there are only two components within the CCCG. (Mosleh et al. 1998).

The advantages of the Beta Factor model, in addition to the ratio model advantages are:

- Simplicity compared to other ratio models.
- Regardless of the number of components comprising the system, it requires the estimation of only two parameters.

The limitations of the Beta Factor model, in addition to the ratio model limitations are:

- The model does not acknowledge CCFs of various multiplicities within the Common Cause Component group. Failure can either be one component or the whole component group. (Hokstad 2004)
- For most redundant systems, this model has been proven to be excessively conservative and pessimistic in predicting CCF failure rates. (Mosleh et al. 1998)

- Rigorous estimators for the beta factor model parameters are fairly difficult to obtain, although approximate methods have been developed and used in practice.(Mosleh et al. 1998) (Mosleh 1986)

### ***1.3.2. C-Factor Model***

The C-Factor model was introduced by Evans et al. (Evans et al. 1984) which is essentially exactly the same as the Beta Factor model with a different interpretation of how the ratio factor should be defined. The C-Factor assumes that the ratio parameter which can be transferred to the specific system under analysis is defined as a ratio of the individual failure rate, not the total failure rate (Hokstad & Rausand 2008):

$$Q_m^{(m)} = C Q_1^{(m)}$$

The C-factor method was developed in an attempt to use the Licensee Event Report (LER) summary data to provide estimates of common cause failure probabilities. The C factor estimator was the fraction of observed root causes of failures that either did, or were judged to have the potential to, result in multiple failures. (Evans et al. 1984)

The advantages and limitations of the C-Factor model are the same as the Beta Factor model.

### ***1.3.3. Multiple Beta Factor (MBF) Model***

Hokstad suggested that the reason for the success of the Beta Model is due to its

extreme simplicity; and to be a success, and proposed a generalized beta model known as the Multiple Beta Factor Model, must be as simple to use in practice (Hokstad 2004) (Hokstad et al. 2006). This simplicity can be achieved by letting the CCF system failure rate for a KooM configuration be calculated as:

$$Q_k^m = C_{KooM} \cdot \beta Q_t$$

$\beta$  = Beta Factor

$C_{KooM}$  = is a configuration factor taking into account the reliability structure of the K-out-of-N system number of components within the CCG

When the configuration factor,  $C_{KooM}$ , equals one, the model is equivalent to the Beta Factor model.

The parameter estimate is (Hokstad 2004):

$$C_{KooM} = \sum_{j=m-k+1}^M \binom{M}{j} G_{j,m} \text{ for } k = 1, 2, \dots, m - 1$$

$$G_{j,m} = \frac{g_{j,m}}{Q\beta} = \sum_{i=0}^{m-j} (-1)^i \binom{m-j}{i} \prod_{x=2}^{j-1+i} \beta_x \text{ for } j = 2, 3, \dots, m$$

$\beta_x$  = The probability that component  $x$  fails given another component has already failed.  $\beta_x = \Pr(A_{x+1} | x \cap \dots \cap A_x)$  where  $A_x$  is a r.v for the event that component  $x$  fails.

$m$  = Common Cause Component Group size

$k$  = Number of failed components due to common cause failure

Further details on the estimation of the MBF parameters are found in (Hokstad 2004).

Advantages over the MGL model is that once the parameters have been estimated the calculation of the basic parameters is greatly simplified. A limitation is that the probability of k-out-of-n failing is independent of the size of n.

#### **1.3.4. Multiple Dependent Failure Fraction (MDFF)**

This Method was proposed by Stamatelatos in 1982 (Stamatelatos 1982) and originally developed for a system of three identical units. This method was modified and extended to four by Hirschberg (Hirschberg 1985).

MDFF is a generalization of the  $\beta$  factor method and can be stated mathematically in the following equation(Anude 1994):

$$\lambda = \lambda_r + \lambda_c = \lambda_r + \lambda n \sum_{n=2}^N f_n = \lambda r + f \lambda$$

$\lambda_r$  = Independent Failure rate component

$\lambda_c$  = CCF rate component =  $\lambda n \sum_{n=2}^N f_n$

$f_n$  = fraction of  $n$  failures ( $n=2, \dots, N$ )

$f$  = fraction of CCFs

Results of the  $\beta$  factor approach are usually more conservative than those of MDFF.

The original articles for this model could not be obtained and so a full assessment could not be completed.

#### **1.3.5. Partial Beta Factor (PBF) Model**

The Partial Beta Factor (PBF) model was first conceived by Edwards in 1982 and later developed by Johnston (Johnston 1987) to allow consideration for the target system dependencies and defenses. After a qualitative analysis identifies CCCGs containing identical components and a criticality assessment of the effect from dependencies based on cut sets, a matrix is created allowing the different attributes leading to dependencies

between the components are evaluated. A Beta Factor is then created as a product of a number of partial beta derived from judgments of system defenses.

$$\hat{\beta} = \prod_j \beta_j$$

$\beta$  = The beta factor for the Beta Factor model.

$\beta_j$  = The partial beta factors for defence  $j$  attribute

The basic parameters are the same as the beta factor model and calculated as:

$$Q_k^{(m)} = \begin{cases} (1 - \beta)Q_t & k = 1 \\ 0 & m > k > 1 \\ \beta Q_t & k = m \end{cases}$$

A limitation of this approach is that each partial beta factor multiplies the whole failure rate for defenses which may only affect a portion of the failure rate. Johnson proposed an extension to the PBF model where the failure rate is separated into specific causes and the Partial Beta Factor only adjusts that portion of the failure rate (Johnston 1987).

$$Q_{m,i}^{(m)} = Q_i \prod_j \beta_{j,i}, \quad Q_m^{(m)} = \sum_i Q_{m,i}^{(m)}$$

$Q_i$  = The failure probability/rate of cause  $i$ .

$\beta_{j,i}$  = The partial beta factor for defence  $j$  and cause  $i$

It should be noted that the outcome of the Partial Beta Factor model is to arrive at a system specific Beta Factor, and as such multiplicity of failures within a common cause component group is not recognized.

### 1.3.6. Multiple Greek Letter Model

Multiple Greek Letter (MGL) Model is a component failure ratio model which the same number of parameters as components within the Common Cause Component Group. This model was introduced by Fleming and Kalinowski (Fleming & Kalinowski 1983) as an extension to the Beta Factor model. The additional parameters were introduced to account for (Anude 1994):

- higher component redundancies,
- failure multiplicities greater than unity, and
- different probabilities of failures for subgroups of the common cause component group.

The basic parameters can be calculated as (Mosleh et al. 1998):

$$Q_k^{(m)} = \frac{1}{\binom{m-1}{k-1}} \left( \prod_{i=1}^k \rho_i \right) (1 - \rho_{k+1}) Q_t$$

$Q_t$  = The total failure probability of one component

$m$  = Common Cause Component Group size

$k$  = Number of failed components due to common cause failure

$\rho_i$  = The MGL parameter.  $\rho_1 = 1, \rho_2 = \beta, \rho_3 = \gamma, \rho_4 = \delta, \dots, \rho_{m-1} = 0$

The MGL estimators are defined as a ratio of the number of components that fail within different configurations. For example  $2n_2$  is the number of components which have failed as part of a CCF event involving two components. This is in contrast with the Alpha Factor Model which defines its parameters in terms of CCF event counts, not component failure counts. For example, the first three parameters of the MGL model are described as:

- $\beta$  the conditional probability that the cause of a component failure will be shared by one or more additional components, given that a specific component has failed.<sup>12</sup>

$$\hat{\beta} = \frac{\sum_{k=2}^m kn_k}{\sum_{k=1}^m kn_k}$$

$n_k$  = number of events involving  $k$  components in a failed state  
 $m$  = the number of components within the CCCG

- $\gamma$  the conditional probability that the cause of a component failure that is shared by one or more components will be shared by two or some additional components, given that two specific components have failed.<sup>13</sup>

$$\hat{\gamma} = \frac{\sum_{k=3}^m kn_k}{\sum_{k=2}^m kn_k}$$

- $\delta$  the conditional probability that the cause of a component failure that is shared by two or more components will be shared by three or more additional components given that three specific components have failed.<sup>14</sup>

$$\hat{\delta} = \frac{\sum_{k=4}^m kn_k}{\sum_{k=3}^m kn_k}$$

A detailed description of the MGL model is provided in NUREG/CR-5485 (Mosleh et al. 1998) including conversions between MGL and AFM.

---

<sup>12</sup> This estimator is for non-staggered testing. (Mosleh et al. 1998)

<sup>13</sup> This estimator is for non-staggered testing. (Mosleh et al. 1998)

<sup>14</sup> This estimator is for non-staggered testing. (Mosleh et al. 1998)

Advantages over the Beta Factor model is that the MGL model can model various multiplicities of failure within the Common Cause Component Group.

### 1.3.7. Alpha Factor Model (AFM)

The alpha factor model (AFM) is a failure event ratio model that was first proposed by Mosleh and Siu in 1987 (Mosleh & Siu 1987). Each  $\alpha_k$  factor is the probability that given a failure it will fail  $k$  components out of  $m$  component within the CCCG. The AFM parameters are defined and calculated as (Mosleh et al. 1998):

$$\alpha_k = \frac{n_k}{\sum_{k=1}^m n_k}$$

$m$  = the number of redundant components

$n_k$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$\alpha_k$  = the fraction of total failure events/frequency that occur in the system resulting in  $k$  out of  $m$  failures.

The alpha factor method is used to estimate the basic event probabilities using (Mosleh et al. 1998) .

$$Q_k^{(m)} = k \binom{m-1}{k-1}^{-1} \cdot \frac{\alpha_k}{\alpha_t} \cdot Q_t \quad \text{non-staggered test data}$$

$$Q_k^{(m)} = \binom{m-1}{k-1}^{-1} \cdot \alpha_k \cdot Q_t \quad \text{staggered test data}$$

where:

$$\binom{m-1}{k-1} = \frac{(m-1)!}{(k-1)! (m-k)!} \quad \text{and} \quad \alpha_t = \sum_{i=1}^m k \alpha_k$$



$$Q_k^{(m)} = \text{basic event failure frequency/probability for } k \text{ components} \\ \text{failing within a common cause component group of size } m, (1 \\ \leq k \leq m).$$

$$Q_t = \text{total failure frequency/probability of each component due to} \\ \text{independent and common cause events.}$$

This formulation has the property that that  $\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_m = 1$  so that the  $\alpha$ s are not mutually independent. (Vaurio 2008).

Due to AFM's ability to calculate its parameters directly from the impact vectors, the AFM is sometimes used as an intermediate step in calculating Beta Factor and MGL parameters. The relationship between these parameters is detailed for different systems within NUREG/CR-5485 (Mosleh et al. 1998).

Advantages over the Beta Factor model is that the AFM model can model various multiplicities of failure within the Common Cause Component Group and unlike the Beta and MGL methods, AFM's parameters are directly related to measurable properties of the system and are usually calculated directly from observed data as impact vectors (Mosleh 1991).

#### **1.4. Shock Models**

Shock models are based on the hypothesis that each component within the CCCG undergoes shocks according to a Poisson process. For each component within the

CCCG the shock is a Bernoulli trial which will fail the component with probability  $\rho$ .

Most shock models are adoptions or simplifications of the multivariate exponential model derived by Marshall and Olkin in 1967 (Marshall & Olkin 1967). For these models the number of failed components,  $k$ , resulting from a shock is binomially distributed. Shock models strive to model the actually physical phenomena that results in CCF to occur.

Shock models have the following advantages:

- Can be used to model high levels of redundancy.(Anude 1994)
- Can estimate CCF frequency even when CCF events have not been observed. (Atwood 1986)
- Easier to adjust for different sizes CCCG groups. (Kvam & Martz 1995)
- Importing/exporting data for different sized systems is more accurate and often easier due to the ability to characterize the underlying probability of common cause failures. (Kvam & Martz 1995)

Shock models have the following disadvantages:

- Includes parameters which are difficult to measure with data (such as shock rates).
- Requires demand/success data to calculate parameters.

- Confusion in the interpretation of single failures being modeled as independent failures, particularly when applying impact mapping rules.
- Lethal shocks need to be distinguished from multiple CCF failing all system components.
- Assumes component symmetry (ie.  $X_{AB} = X_{BC}$ ).
- Assumes that given a shock has occurred, items will fail independently which may be violated in practice. (Anude 1994)
- Assumes zero time to repair. (Atwood 1986)
- Assumes renewal to as good as new. (Atwood 1986)
- Assumes Constant Failure Rates. (Atwood 1986)
- The  $\rho$  parameter is independent of the size of the CCF. (Vaurio 1999)
- Any subset of  $k$  components of a system of size  $m$  is equally vulnerable to exactly the same common-causes and stresses as in a system of size  $k$ , or anything larger than  $k$ . This results in the assumption that  $n_k > n_{k+1}$  (the mapping rule). (Vaurio 1999)
- Data is needed from a system with at least  $m=3$  in order to solve the three unknowns. (Vaurio 2008)
- The analyst needs to distinguish between a single CCF and a single independent failure. This can become subjective from fault reports leading to higher uncertainty.

- No parameters are directly linked to the degree of system protection against CCFs.
- Probability that the value of the binomial parameter  $\rho$  remains fixed across all system shocks despite shocks having different intensities and different sources. (Anude 1994)
- Does not model different intensity shocks to the system. (Anude 1994)
- Parameter calculation can be cumbersome. (Kvam 1993)
- No inference can be made given knowledge of the failure cause.
- The model cannot account for unique system architecture which may contribute or defend against CCF.

#### ***1.4.1. Binomial Failure Rate Model***

The Binomial Failure Rate Model (BFRM) model was proposed by Vesely in 1977 (Vesely 1977) to adapt the shock model proposed by Marshall and Olkin. This model was motivated by estimation with less data than previously required and to describe the underlying failure process generated by CCF events. It assumes that CCF occur when all  $m$  redundant components of a system are challenged by a shock at a rate of  $\mu$ . The number of resulting failures from each shock,  $k$ , is random with a binomial distribution with probability  $\rho$ .

This model has also been known as the three-parameter BFR model with parameters,  $Q_I$  (or  $\lambda$ ),  $\mu$  and  $\rho$ . These parameters can be estimated using (Marshall et al. 1998):

$$Q_I = \frac{n_I}{mN_D}$$

$$\sum_{k=1}^m kn_k = \rho \frac{m \cdot n_t}{1 - (1 - \rho)^m} \quad \text{solve for } \rho$$

$$\mu = \frac{n_t}{N_D} \cdot \frac{1}{1 - (1 - \rho)^m}$$

where

$$n_t = \sum_{i=1}^m n_k$$

$n_k$  = the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$n_I$  = the number of failure events/frequency which resulted in the independent failure of the component.

$n_t$  = total number of common cause failures.

$N_D$  = the number of demands on the system (or time  $T$ ), can also be called  $N_S$

The basic parameters can be calculated as (Vesely 1977):

$$Q_k^m = \begin{cases} Q_I + \mu \cdot \rho(1 - \rho)^{m-1} & \text{where } k = 1 \\ \mu \cdot \rho^k(1 - \rho)^{m-k} & \text{where } 2 \leq k \leq m \end{cases}$$

$Q_I$  = the independent failure rate of each component

$Q_k^m$  = basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$\mu$  = rate of shocks

$\rho$  = probability of component failure given a shock

The rate of failure of  $k$  components is simply the binomial probability of  $k$  in  $m$  components failing multiplied by the rate of shocks. The rate of failure for a single

component is the independent failure rate plus the contribution of 1 component failing due to a common cause shock. Probability that the value of the binomial parameter  $\rho$  remains fixed across all system shocks.

Due to its inaccuracy to real systems, the model presented here is rarely used (Mosleh et al. 1988) (Kvam 1998b); instead, a simple binomial shock model using the BFR model with lethal shocks is typically preferred. Note that the BFRM and  $\beta$ -factor model are the same for a two component system. (Rausand & Høyland 2003)

#### ***1.4.2. Binomial with Lethal Shocks***

Atwood proposed an extension to the BFR model in 1986 that included an additional independent process of lethal shocks (Atwood 1986). In this model, each lethal shock will fail all components of the system at a rate of  $\omega$ .

This model has also been known as the four-parameter BFR model with parameters,  $Q_I$  (or  $\lambda$ ),  $\mu$ ,  $\rho$  and  $\omega$ . These parameters can be calculated using [85]:

$$Q_I = \frac{n_I}{mN_D}$$

$$\sum_{k=1}^m kn_k = \rho \frac{m \cdot n_t}{1 - (1 - \rho)^m} \quad \text{solve for } \rho$$

$$\mu = \frac{n_t}{N_D} \cdot \frac{1}{1 - (1 - \rho)^m}$$

$$\omega = \frac{n_L}{N_D}$$

where

$$n_t = \sum_{i=1}^m n_k$$

$n_k =$  the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$n_1 =$  the number of failure events/frequency which resulted in the independent failure of the component.

$n_L =$  total number of lethal common cause failures.

$n_t =$  total number of common cause failures.

$N_D =$  the number of demands on the system (or time  $T$ ), can also be called  $N_S$

The basic parameters can be calculated as (Atwood 1986):

$$Q_k^m = \begin{cases} Q_1 + \mu \cdot \rho(1 - \rho)^{m-1} & \text{where } k = 1 \\ \mu \cdot \rho^k(1 - \rho)^{m-k} & \text{where } 2 \leq k < m \\ \mu \cdot \rho^m + \omega & \text{where } k = m \end{cases}$$

$Q_1 =$  the independent failure rate of each component

$Q_k^m =$  basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$\mu =$  rate of shocks

$\rho =$  probability of component failure given a shock

$\omega =$  rate of lethal shocks

This extension has been found to be more accurate than the basic BFR model (Mosleh et al. 1988). The probability that the value of the binomial parameter  $\rho$  remains fixed across all system shocks.

### 1.4.3. Rho Distribution Models

The major limiting assumption of the BFRM is that for each shock, the probability of  $k$  components failing,  $\rho$ , is constant. This is improbable as each shock to the system is likely to have a different intensity. Three very similar models modeled  $\rho$  to be a random variable with a beta distribution:

- The Random Probability Shock model as suggested by Hokstad in 1988. (Hokstad 1988).
- The Distributed Failure Probability model was proposed by Hughes in 1986. (Hughes 1987)
- The BFR Mixture Model was proposed by Kvam in 1998 (Kvam 1998b) (Vaurio 1999). An adaption was made to allow a non-parametric distribution for  $\rho$ . (Kvam 1998a)

These models provides the ability to include various degrees of dependence between components through the combination of the  $\beta$ -Factor model and the BFR model. The  $\beta$ -factor and BFR models are special cases of the RPS model.

The model replaces the BFRM fixed parameter  $\rho$  with a beta distribution:

$$\rho \sim B(r, s)$$

In the RPS model the beta distribution parameters  $r$  and  $s$  are transformed into more meaningful parameters  $Q$  and  $D$ :

$$Q = \frac{r}{r + s} \quad \text{and} \quad D = \frac{1}{r + s + 1}$$



Q is the mean of the beta distribution and therefore can be defined as the point estimate of  $\rho$ . D can be related to the variance of the beta distribution and is considered to be a measure of dependence on the outcomes of the shocks to various components. (Zitrou 2006b)

#### ***1.4.4. Multinomial Failure Rate Model***

The Multinomial Failure Rate Model (MFR) was proposed by Apostolakis and Moieni in 1987 (Apostolakis & Moieni 1987). This is a shock model which prescribes the condition  $\phi_1 + \phi_2 + \dots + \phi_m = 1$ . Where  $\phi_k$  is the conditional fraction of exactly k failures in the event of a system shock (Anude 1994).

The original articles for this model could not be obtained and so a full assessment could not be completed.

#### ***1.4.5. Stochastic Reliability Analysis Models***

The Stochastic Reliability Analysis (SRA) Models were proposed by Dörre in 1989 (Dörre 1989). This model differs in approach to others by assuming that dependent failure is the basic phenomenon, while independent failure refers to a special limiting case. This results in a shock model which replaces  $\rho$  with any distribution  $g(\rho)$  in a more general case to the BFR Mixture Model.

This model recognizes that the total failure rate for a component is simply the sum of the failure rates for different causes. CCF event modeling is treated as a special cause where the failure causes are shared. The model was criticized over confusion regarding definition of the causes and environments (Parry 1989).

#### ***1.4.6. Trinomial Failure Rate Model***

The Trinomial Failure Rate Model (TFR) was proposed by Han et al. in 1989 (Han et al. 1989). This is shock model which amends the binary states of ‘working’ or ‘fail’ to include a ‘grey’ condition. This grey condition includes partial failures, incipient failures and potential failures. The probability of a component state is split between  $p$ ,  $q$  and  $r$  for failed, grey and operating state respectively, given a CCF shock occurs to the system.

Models which use the impact vector methodology, the ‘grey condition’ is accounted for through methods which deal with uncertainty.

#### ***1.4.7. Multi-Class Binomial Failure Rate Model***

The Multi-Class Binomial Failure Rate (MCBFR) model was first proposed by Hauptmanns in 1996 (Hauptmanns 1996). This model attempts to maximize the information available from the collected CCF data to increase the accuracy of the BFR model. It achieves this by assigning observations to different classes according to their technical characteristics and applying the BFR formulation to each of these classes.

The results are determined by a superposition of BFR expressions for each class with a coupling factor.

This model attempts to determine the actual shock rate of the failure modes by extracting data from other systems and applying this through a coupling factor to the system being analyzed. The model increases the sources of data which can be used to analyse a system from other systems subject to the same failure mechanism. This is at the cost of requiring more information from the data and introducing some subjectivity.

$$Q_k^m = \begin{cases} Q_I + v \cdot \sum_{l=1}^L h_l \rho_l (1 - \rho_l)^{m-1} & \text{where } k = 1 \\ v \cdot \sum_{l=1}^L h_l \rho_l^k (1 - \rho_l)^{m-k} & \text{where } 2 \leq k \leq m \end{cases}$$

where:

$$\sum_{l=1}^L h_l = 1, \quad l = 1, \dots, L, \quad v = \mu + \omega$$

$Q_I$  = the independent failure rate of each component

$Q_k^m$  = basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ).

$\mu$  = rate of shocks

$\omega$  = rate of lethal shocks

$v$  = total shock rate

$\rho_l$  = probability of component failure via mechanism  $l$ , given a shock

$h_l$  = conditional probability that a shock will cause component failures via mechanism  $l$ .

The lethal shock rate is incorporated into the total shocks to the system because the case where all components fail will be handled by the separation of the failure

mechanisms and the coupling factor. The parameter estimates for this model are detailed in (Hauptmanns 1996).

While the model adjusts the fragility parameter based on the cause/mechanism the shock rate remains a single rate/probability. Does not model different intensity shocks to the system. Parameter calculation can be cumbersome.

#### ***1.4.8. The Coupling Model***

The Coupling Model was first proposed by Kreuser and Peschke in 1997 (Kreuser & Peschke 1997) (Kreuser & Peschke 2001). It is an extension of the BFR driven by the necessity to capture the uncertainty in the interpretation of data from each CCF event when applied to the system of interest. The Coupling Model captures two additional sources of uncertainty; translation uncertainty and interpretation uncertainty.

Translation uncertainty describes the uncertainty of CCF data coming from various sources from the system of interest, and is captured by a new parameter called the Applicability Factor,  $f$ .

Interpretation uncertainty stems from the classification of the component failure state across particular classes (eg. failed, degraded, incipient) which can be unclear or missing in collected data. The interpretation uncertainty is captured in a mixture of beta distributions which capture the uncertainty of the parameter  $p$  from the Binomial

Failure Rate model which in this model is called the Coupling Factor. The approach to weigh the interpretation uncertainty between the alternative outcomes is similar in approach to the treatment of interpretation uncertainty used in creating impact vectors discussed in Chapter 2.

The basic parameters can be calculated as (Kreuser & Peschke 2001):

$$Q_k^{(m)} = \sum_{j=1}^N P_{j,k/m}$$

$$P_{j,k/m} = \frac{T_{CCF_j} \cdot f_j}{T_{obs}} \binom{m}{k} p_j^k (1 - p_j)^{m-k}$$

$T_{CCF}$  = failure detection time (test interval)

$T_{obs}$  = the total observation time.

$f_j$  = applicability factor

$p_j$  = coupling factor

The advantage of the Coupling model is that it calculates the coupling strength based on the observed phenomenon for each CCF event and it includes the mechanisms to quantify uncertainties in expert judgments and observed data when estimating CCF probabilities (Kreuser & Peschke 2001).

#### **1.4.9. Bayes Testing and Estimation BFR Model**

The Bayes Testing and Estimation BFR Model was proposed by Kvam and Martz in 1995 (Kvam & Martz 1995). This model recognized that most BFR models need to distinguish between independent failures and CCF from system shocks, which can be

difficult because a single failure could have been independent or the result that only one component failure from a system shock. Likewise two components failing could be a coincidence of two independent failures or two components failing from a system shock. This model does not require this distinction to be made.

This model attempts to limit the number of parameters as opposed to most BFR extensions due to increased difficulty in estimating parameters. This is done by considering CCF as lethal shocks only. The parameter  $\rho$  has been modeled as a random variable with a beta distribution.

This model is restricted to systems of low redundancy due to the assumption all CCF are lethal.

### ***1.5. Interference Models***

Interference models also attempt to model the physical phenomena of CCF but without the shock model's assumption of independence. Instead these models predict the number of failures by assuming random variables for component strength and load. When the load exceeds the strength a component is expected to fail. The more intense the load or the more depleted the strength than the higher the probability of failure. There is no explicit distinction between an independent and common cause failure event.

Inference models have the following advantages:

- Can be used to model high levels of redundancy.
- Can estimate CCF frequency even when CCF events have not been observed.
- Easier to adjust for different sizes CCG groups.
- Does model different intensities of shocks to the system.
- Directly models the system's protection against CCF through the resistance measure.
- There is no need to distinguish between a single CCF and a single independent failure.
- Lethal shocks are quantified by their shock intensity and included within the model formulation.
- Importing/exporting data for different sized systems is more accurate and often easier due to the ability to characterize the underlying probability of common cause failures. (Kvam & Martz 1995)

Interference models have the following disadvantages:

- Requires a probability distribution to be estimated for shock and resistance intensities. This requires knowledge of the physical characteristics of the components and the data required to quantify distributions differs from just failure and success data.
- Requires demand/success data to calculate parameters.

- Assumes component symmetry (ie.  $X_{AB} = X_{BC}$ ).
- Assumes zero time to repair.
- Assumes renewal to as good as new.
- Assumes Constant Failure Rates.
- Any subset of  $k$  components of a system of size  $m$  is equally vulnerable to exactly the same common-causes and stresses as in a system of size  $k$ , or anything larger than  $k$ . This results in the assumption that  $n_k > n_{k+1}$  (the mapping rule). (Vaurio 1999)
- No inference can be made given knowledge of the failure cause.
- The model does not explicitly account for unique system architecture which may contribute or defend against dependencies between components.

### ***1.5.1. Common Load Model***

The Common Load model proposed by Mankamo and Kosonen in 1977 (Mankamo 1977) is based on a load-strength interference methodology for describing the failure mechanism. The model interprets the failure mechanism as a load imposed on a component where the components strength is tested. A failure occurs when the resistance is not sufficient to withstand the load.

When it comes to redundant systems of components, the load posed to the system is shared by all the components of the system equally, and a failure of certain multiplicity is determined by the number of components whose resistance is exceeded by the load.



Both the load and the component resistance are described in terms of random variables and assumed probability distributions (Zitrou 2006b).

The probability density function of the resistance,  $R$ , is denoted by  $f_R(x)$ . In the event of an occurrence of a random shock,  $S$ , with a probability density function of  $g_s(x)$ , then the event of having exactly  $k$  of the components fail simultaneously, is given as (Anude 1994):

$$Q_k^{(m)} = P(R_k \leq S < R_{k+1})$$

$$Q_k^{(m)} = \int_0^{\infty} \frac{m!}{k!(m-k)!} (F_R(y))^k (1 - F_R(y))^{m-k} g_s(y) dy$$

$S$  = the random variable for the shock intensity

$R_k$  = the random variable for resistance intensity where  
 $R_1 \leq R_2 \leq \dots \leq R_n$

$g_s(x)$  = the probability distribution for the shock random variable

$F_R(x)$  = the cumulative probability distribution for the resistance random variable

$k$  = the multiplicity of failure being investigated

$m$  = the number of components within the CCG

The model has a fixed number of parameters, independent of the size of the system. Like the Shock Models, the model can be applied to any failure multiplicities. The model assumes that the  $n$  components of a system have independent and identically distributed random resistances  $R_1, R_2 \dots R_n$ . (Anude 1994)

Cases of non-symmetry can be modeled by removing the assumption of identical distributed components and creating separate  $f_R(x)$  distributions for each component.

### ***1.5.2. Inverse Stress-Strength Interference Model (ISSI)***

The Inverse Stress Strength Interface Model (ISSI) was proposed by Guey in 1984 (Guey 1984) to minimize the data requirements when compared to the Common Load Model. By inverting the expressions for the interference models, a relationship between the model parameters and the known failure rates and dependencies shown in the data. The specific formulas depend on the distributions assumed for the resistance and shock PDFs.

The ISSI methods, discussed in detail by Guey, demonstrate a series of assumptions and techniques to estimate an interference model using different levels of available evidence such as failure rates and laboratory tests.

### ***1.5.3. Harris Model***

The Harris Model was proposed by Harris in 1986 as an extension to the Common Load Model (Harris 1986) to allow for different mission times and partial failures.

Let  $N(t)$  be the number of shocks arriving at or before time  $t$ , where  $0 \leq t \leq T$ . If  $N(T)$  shocks have arrived in  $[0, T]$ , the arrival times are designated by  $0 < t_1 < \dots < t_n < T$ . The shocks have random, independent and identically distributed magnitudes,  $X(t_1), X(t_2), \dots, X(t_n)$ . Each component has a resistance magnitude with random variable  $Y_1, Y_2, \dots, Y_m$ . Component  $i$  fails at time  $t_j$  when  $X(t_j) > Y_i$ . Then  $k$  components fail at time  $t_j$  wherever  $Y_k < X(t_j) \leq Y_{k+1}$ . (Anude 1994)

In order to model the degraded state of components, the additional function  $H(x, y)$  is introduced (Anude 1994):

$$Y_i(t_j^*) = H(Y_i(t_j), X(t_j))$$

The original article for this model could not be obtained and so a full assessment could not be completed.

#### ***1.5.4. Knowledge Based Multi-dimension CCF Model (KBMD)***

The Knowledge Based Multi-Dimension CCF Model (KBMD) was proposed by Liyang Xie in 1998 (Xie 1998). The KBMD model presents the ‘root cause’ and ‘coupling mechanism’ as a random variable representing the environment load which provides dependencies. This model recognizes that the system will be exposed to many different environmental ‘shocks’; therefore, the model superimposes these load-strength relationships through multi-dimension environment load-component strength interference analysis. The KBMD model uses a discretization of the continuous model in order to simplify this complex multi load strength superposition.

This model also accounts for multiple failure events occurring between inspection intervals by comparing real CCF failure (instantaneous detection) to the relative CCF (detection at inspection intervals).

## 1.6. Other Models

### 1.6.1. *Square Root Bounding Method*

The Square Root Bounding Method was developed by the United States Reactor Safety Study: WASH 1400 (Rasmussen 1975), and further refined by Martin and Wright in 1987 (Martin & Wright 1987). The method has been criticized for lacking practical foundations. It makes two major assumptions, that the failure probability of the redundant system cannot be higher than the failure probability of a component, and the system failure probability cannot be lower than the system failure probability without CCF (Anude 1994).

$$P_s \leq P_C \leq P_i$$

$P_C$  =failure probability of system with CCF  
 $P_s$  =failure probability of system without CCF  
 $P_i$  =failure probability of a single item

One further assumption is that the value of  $P_C$  follows a log-normal distribution symmetrical about  $P_s$  and  $P_i$ . The median to this distribution is found to be (Anude 1994):

$$P_C = \sqrt{P_s \cdot P_i}$$

Despite being one of the first CCF models used, it is no longer used.

### 1.6.2. *Implicit Method*

The Implicit Method was introduced by Fleming, K.N. and Mosleh, A. in 1985 (Fleming & Mosleh 1985). In this method, the fault trees are built without considering CCF and then the algebraic system unreliability expression is derived. The expression

is then evaluated to ensure that the contribution of CCF is correctly included. This procedure is commonly used with many of the models as a qualitative means to identify dominant CCCGs.

Implicit methods are inappropriate to the CCF analysis for large-scale fault trees. This method, usually the algebraic expression of system unreliability is not easy to derive. This makes the implicit method feasible only in hand-calculating the unreliability of relatively small systems. Implicit methods consider CCF in the process of analysis rather than in modeling stage; thus, they do not need to perform the monotonous basic event expansion to include CCF. (Vaurio 1998)

### ***1.6.3. Reliability Cut Off Method***

The Reliability Cut Off Methods was proposed by Bourne et. al in 1981 (Bourne et al. 1981) as a system level estimate of CCFs based on an assessment of the vulnerability of the system. No identification of CCCG is conducted and the methodology assumes that the unreliability of a system due to CCFs can never exceed some limiting values, determined by system design. These estimates do not involve the use of data and instead use generic estimates from experts.

The original article for this model could not be obtained and so a full assessment could not be completed.

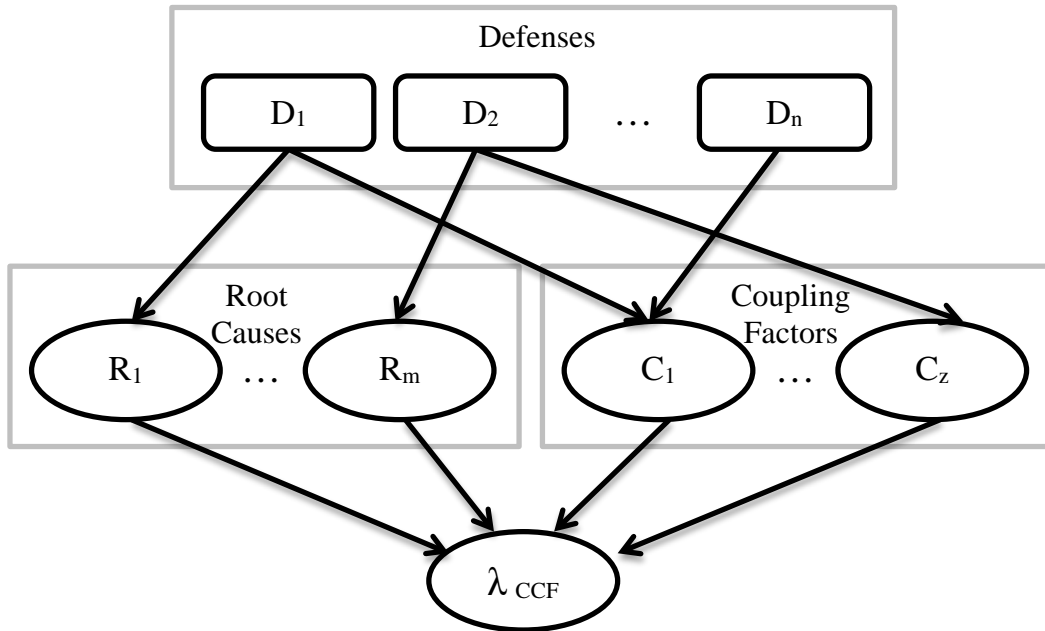
#### ***1.6.4. Unified Partial Method***

The Unified Partial Method (UPM) (Brand & Gabbot 1993) is the current method which has been adopted by the UK nuclear industry. UPM is a methodology to assess the vulnerability of a system to CCF and uses one of two models to quantify its estimates, the Partial Beta Factor method for component level analysis, and the Cut-Off method for system level analysis. Brand describes UPM as not being a complete method for dependent failure assessment, but a useful methodology for ‘standard systems’ (Mosleh et al. 1998).

#### ***1.6.5. Influence Diagram Model (Zitrou 2006a)***

Zitrou in 2006 proposed an extension of the UPM model using influence diagrams and a more detailed mathematic formulation using Bayesian methods. The objective of Zitrou’s research was to explore the modeling of CCF using advanced mathematical techniques (influence diagrams). Zitrou wanted to keep the desirable features of UPM where attributes of the system are included in the model, the ability to provide estimates in the absence of data and the simplistic application of the method by analysts. Zitrou wanted to use the influence diagram to extend UPM’s accuracy by modeling the dependency between defenses and improve the models quantitative estimates.

Zitrou’s model consisted of the creating of an influence diagram which in general terms followed the convention of figure 5. (Zitrou 2006a, p.18)



**Figure 65: Zitrou General Influence Diagram Structure**

The specific taxonomy used to define the ID nodes were the same as for UPM. The specific dependencies between nodes were established using an expert elicitation technique.

Two unique elements are proposed in Zitrou's model (Zitrou 2006a, p.257).

- The definitions of the dependencies between defenses were established to determine if improving one defense would have a positive, negative or natural effect on another defense.
- A geometric scaling model was proposed which is used to quantify the effect of the defense levels on the probability of root causes and coupling factors. This model reduces the burden of the quantification process by allowing the root cause and coupling factor probability distributions to be determined based on a base defense level. The geometric scaling model can then scale the probability distributions dependent on the level of defense applicable.

Zitrou's model achieves the following objectives (Zitrou 2006a, p.254):

- Incorporates the qualitative advantages of the UPM model.
- After quantification by experts the model can be easily used by practitioners.
- Extends the casual modeling of UPM to a finer level.
- Captures the dependency between defenses.
- Captures the uncertainty of the expert judgment.
- Provides an investigative framework in which conditional probabilities can be explored.

Zitrou's research was to conduct an exploration of using influence diagrams to model CCF. The thesis proposed a methodology and conducted the limited development of a quantification model on Emergency Diesel Generators using some expert estimation. The example model was not fully developed and verification against known system results was not conducted. The model did not consider incorporating data analysis techniques from CCF databases.

#### ***1.6.6. CCF model for Event Assessment (Kelly et al. 2011)***

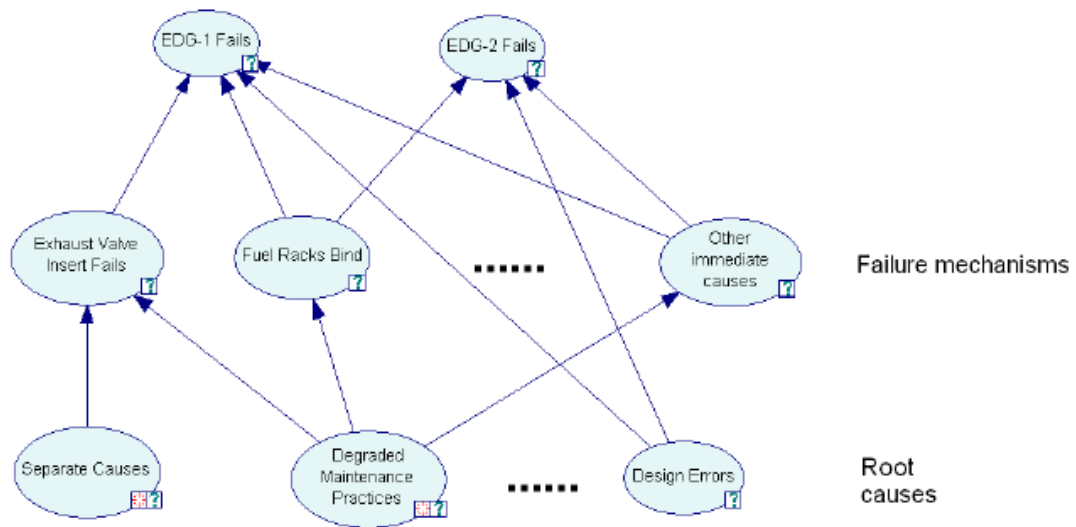
Kelly et al., have written a draft paper which demonstrates the inability of current models to conduct event assessments. The paper proposes using a Bayesian Network to model the causal relationship between root causes, failure mechanisms and the CCF event probability. Two methods of constructing this model are proposed:

- The first model explicitly models the root causes and failure mechanisms specific to the component.



- The second model uses the generic CCF taxonomy used by the INL CCF database.

The paper focuses on conducting event assessments where a failure cause is known; as such the model does not include the coupling factors or defenses. An example of such a model is included as Figure 66. This paper expresses an ideology for CCF modeling but does not propose specific model construction or quantification details. This paper forms the objective of this research.



**Figure 66: Bayesian network representing more general situation of multiple failure mechanisms and causes in a CCG of two EDGs (Kelly et al. 2011, p.6)**

### 1.6.7. Physics-Based CCF

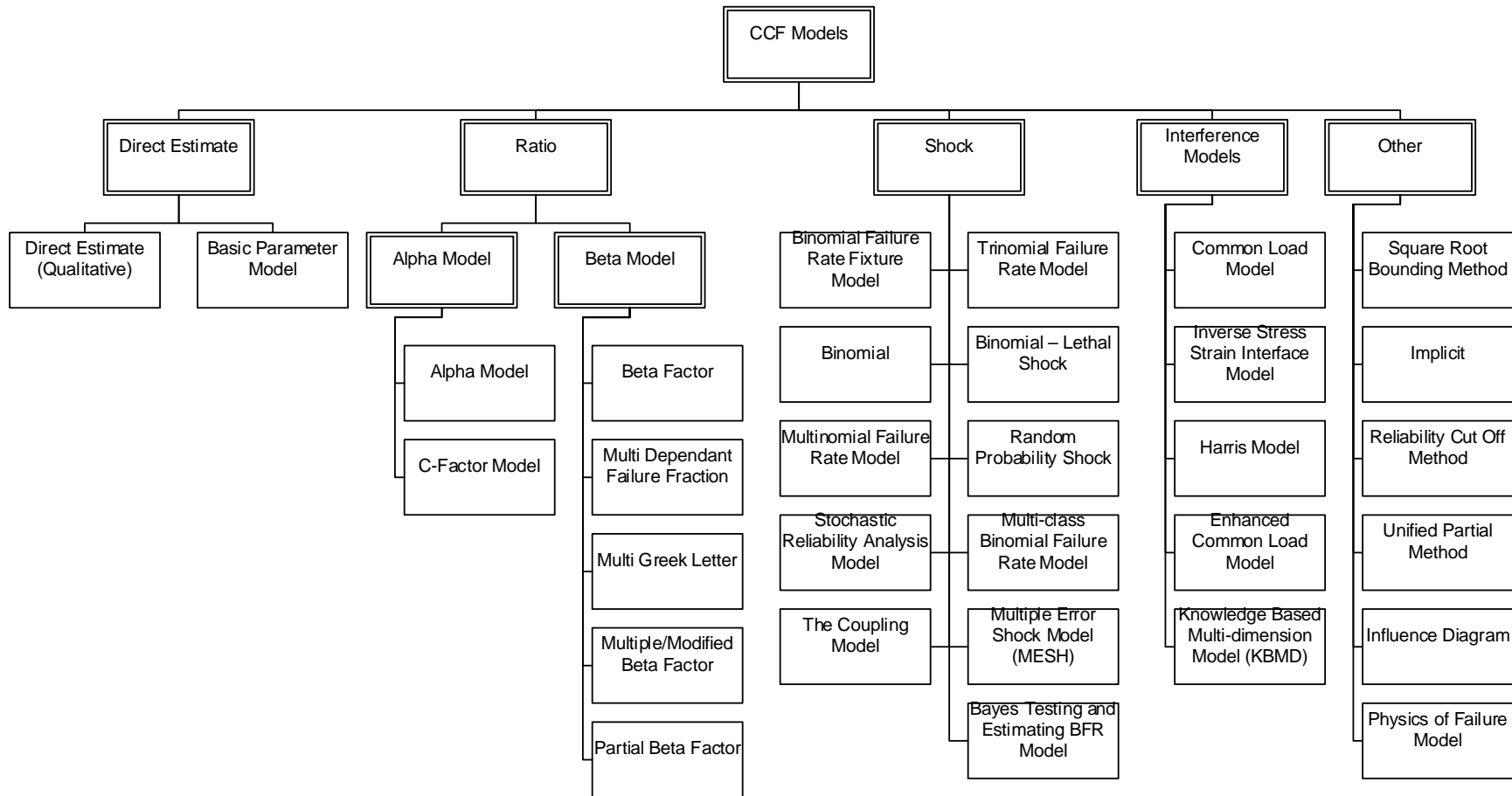
In 2011, Mohaghegh et al., proposed a model to explicitly model the failure mechanisms which may couple components (Mohaghegh et al. 2011). This involves

modeling the individual failure mechanisms such as fatigue and wear, applying a mechanistic approach to determine the interaction of the failure mechanism between components, and then integrating the physics of failure model into the PRA. The paper provides a theoretical foundation for using such a model. The methods current limitations are acknowledged such as expanding the physics of failure models from the material level to the component level and the lack of physics of failure modeling for many of the required failure mechanisms.

This model has further challenges such as the move from implicit to explicit modeling which puts into question whether the failures could be classified as CCF (see chapter 3). CCF models account for the known-unknowns. We know unexpected failures will occur that will fail multiple components, but we don't know the failure mechanism. Therefore such a model is unlikely to be able to model unexpected shocks such as a person positioning a ladder such that it fails two EDGs or an outbreak of jelly fish clogging water intakes. Despite these challenges, the proposed model proposed an excellent foundation to move failures from CCF models to explicit modeling within PRAs.

### ***1.7. Model Comparison***

Each model has been categorized based on the features of the model and its assumptions. The summary of model classification is shown in Figure 67. The features of each model is compared in Table 49.



**Figure 67: Classification of CCF Models**

**Table 49: Comparison of CCF model features**

	General Dependency Model	Partial Alpha Factor Model	Basic Parameter	Beta Factor	Partial Beta Factor	Alpha Factor Model	Binomial Failure Rate Model with Lethal Shocks	Common Load	Reliability Cut Off	Influence Diagram	Bayesian Network
Feature Description	GDM	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Explicitly Models System Features	GDM	PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Models failure cause	Y	Y	N	N	P	N	N	N	P	Y	Y
Models failure cause defense	P	N	N	N	Y	N	N	N	Y	N	P
Models coupling factor	Y	Y	N	N	P	N	N	N	P	Y	N
Models coupling factor defense	P	P	N	N	Y	N	N	N	Y	N	N
Models deeper causal levels	Y	N	N	N	N	N	N	N	N	N	Y
Models cause condition / shock	Y	N	N	N	N	N	Y	Y	N	N	Y
Models multiplicity of failures within CCCG	Y	Y	Y	N	N	Y	Y	Y	N	N	Y
Models includes consideration for rectification period	N	N	N	N	N	N	N	N	N	N	N
Common Cause Component Grouping Characteristics		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Model non-symmetrical but similar components within the same CCCG	Y	Y	N	N	N	N	N	N	N	N	Y
Model different components within the same CCCG	Y	N	N	N	N	N	N	N	N	N	Y
A component can be part of many CCCGs	Y	Y	N	N	N	N	N	N	N	N	Y
No limit to CCCG size	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Model different failure multiplicities within the CCCG ( $k$	Y	Y	Y	N	N	Y	Y	Y	N	N	Y

Event Assessment Capabilities		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Event Assessment with knowledge of a failed component	Y	Y	Y	N	N	Y	Y	?	N	Y	Y
Event Assessment with knowledge of failure cause	Y	Y	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence - Partial Failures	Y	N	N	N	N	N	N	N	N	Y	Y
Uncertain Evidence- Virtual evidence of cause	Y	N	N	N	N	N	N	N	N	Y	Y
Parameter Estimation		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Impact Vector Method (including method for incorporating	P	Y	Y	P	N	Y	Y	N	N	N	N
Expert estimations (in absence of any data)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Account for reliability growth (discount previous failures)	N	N	N	N	N	N	N	N	N	N	N
Update parameters with new evidence	Y	Y	Y	P	N	Y	Y	Y	N	N	N
Incorporate evidence from different sized CCCGs	P	Y	N	P	N	P	Y	Y	N	N	N
Account for CCF which occurred in a different mission time	N	N	N	N	N	N	N	N	N	N	N
Account for CCF data which has artificial separation in time due	N	N	N	N	N	N	N	N	N	N	N
Use system specific failure rate data combined with generic	Y	Y	N	Y	N	Y	N	N	N	N	N
Uncertainty Characteristics for Parameter Estimation		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Does not require distinguish between independent and single	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y
Failures outside the mission period	Y	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty of shared cause	Y	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty of coupling factor	Y	Y	Y	P	N	Y	Y	N	N	N	N
Uncertainty in intervals due to staggered testing	P	P	P	P	N	P	P	N	N	N	N
Partial failures and component degradation	Y	Y	Y	P	N	Y	Y	N	N	N	N
Usability and Cultural Considerations		PAFM	BP	BF	PBF	AFM	BFRL	CL	RCO	ID	BN
Backward compatible to Alpha Factor Model parameters	Y	Y	Y	N	N	Y	N	N	N	N	N
The time investment is no more than the alpha factor model.	P	Y	Y	Y	Y	Y	Y	N	Y	N	N
Automatic parameter estimation is possible from the	P	Y	Y	Y	N	Y	Y	N	N	N	N

## Appendices 2: Detailed Description of Existing Failure Data

### Taxonomy

#### **2.1. Introduction**

The following are verbatim definitions for the CCF classification taxonomy contained in NUREG/CR-6268 Rev 1 (Wierman et al. 2007). This reproduction is to aid in the understanding of chapter 4.

#### **2.2. Failure Causes:**

##### **Design/construction/manufacture Inadequacy.**

Encompasses actions and decisions taken during design, manufacture, or installation of components both before and after the plant is operational.

##### **Operations/Human Error (Plant staff error)**

Represents causes related to errors of omission and commission on the part of plant staff. This category includes accidental actions and failure to follow procedures for construction, modification, operation, maintenance, calibration, and testing. It also include ambiguity, incompleteness or error in procedures for operation and maintenance of equipment. This includes inadequacy in construction, modification, administrative, operational, maintenance, test and calibration procedures.

**External Environment:**

Represents causes related to a harsh external environment that is not within component design specifications. Specific mechanisms include electromagnetic interference, fire/smoke, impact loads, moisture (sprays/floods etc) radiation, abnormally high or low temperature, and acts of nature.

**Internal to Component:**

Associated with the malfunctioning of something internal to the component. Internal causes result from phenomena such as normal wear or other intrinsic failure mechanisms. It includes the influence of the internal environment of a component. Specific mechanisms include erosion/corrosion vibration, internal contamination, fatigue and wear out/end of life.

**State of other component:**

The component is functionally unavailable because of failure of a supporting component or system. CCF events exclude those events that have dependencies that would reasonably be expected to be modeled in an individual plant examination or PRA

**Unknown:**

Used when the cause of the component state cannot be identified.

**Other:**

Used when the cause cannot be attributed to any of the previous cause categories. This category is most frequently used for cases of setpoint drift.

**2.3. Coupling Factor Definitions****2.3.1. *Environmental Based******Environmental External:***

Refers to all redundant systems/components exposed to the same external environmental stresses (e.g flood, fire, high humidity and earthquake). The impact of several of these environmental stresses is normally modeled explicitly in current PRAs (by analyzing the phenomena involved and incorporating their impact into the plant/system models.) Other environmental causes such as high humidity and temperature fluctuations are typically considered in CCF analysis and treated parametrically.

***Environment Internal:***

Refers to commonality of multiple components in terms of the medium of their operation such as internal fluids (water, lube oil, gas etc)



### **2.3.2. Design Based**

#### ***Hardware design system:***

System-level coupling factors include features of the system or groups of components external to the components that can cause propagation of failures to multiple components.

#### ***Hardware design Parts:***

Component-level coupling factors represents features within the boundary of each component.

### **2.3.3. Quality Based**

#### ***Quality Install:***

Covers both initial and later modifications and refers to the same construction/installation staff, construction/installation procedure, construction/installation testing/verification procedure and the construction/installation schedule.

#### ***Quality manufacturing:***

Refers to the same manufacturing staff, quality control procedure, manufacturing method and material.

#### **2.3.4. Maintenance Based**

##### ***Operations Maintenance Schedule:***

Maintenance/Test/Calibration schedule refers to the Maintenance/Test/Calibration activities on multiple components being performed simultaneously or sequentially during the same event.

##### ***Operations Maintenance Procedure:***

Refers to propagation of errors through procedural errors and operator interpretation of procedural steps. It is recognized that for non-diverse equipment, it is impractical to develop and implement diverse procedures.

##### ***Operations Maintenance Staff:***

Refers to the same maintenance/test/calibration team being in charge of maintaining multiple systems/components

#### **2.3.5. Operation Based**

##### ***Operations Operational Procedure:***

Refers to the cases when operation of all (functionally or physically) identical components is governed by the same operating procedures. Consequently, any deficiency in the procedures could affect these components. Sometimes, a set of procedures or combination of procedure and human action act as the proximate cause and coupling factor. In other cases, a common procedure results in failure or multiple

failures of multiple trains.

***Operations Operational Staff:***

Refers to the events that result in the same operator (team of operators) is assigned to operate all trains of a system, increasing the probability that operator errors will affect multiple components simultaneously.

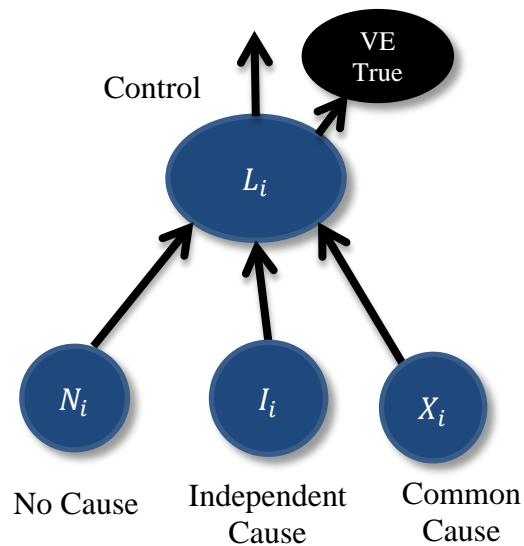
## Appendices 3: Calculation of Mutually Exclusive Nodes

### Using Control Node

The following appendix shows algebraically that the probability of each state in the control node is equal to the mutually exclusive parent node probabilities.

#### 3.1. The Bayesian Network

The Bayesian Network which provides modeling of the cause conditions is shown in Figure 68.



**Figure 68: Structure of mutually exclusive Bayesian network with VE**

The Conditional Probability Tables for the three parent nodes,  $X_i, I_i, N_i$  are:

**Table 50: CPT for Common Cause Condition  $X_i$**

$X_i$ State	
$X_i$	$\eta_i Q_{E,i}$
$\bar{X}_i$	$1 - \eta_i Q_{E,i}$

**Table 51: CPT for Independent Cause Condition  $I_i$**

$I_i$ State	
$I_i$	$(1 - \eta_i) Q_{E,i}$
$\bar{I}_i$	$1 - (1 - \eta_i) Q_{E,i}$

**Table 52: CPT for No Cause Condition  $N_i$**

$N_i$ State	
$N_i$	$1 - Q_{E,i}$
$\bar{N}_i$	$Q_{E,i}$

The Conditional Probability Table for the control node,  $L_i$  is:

**Table 53: CPT for Control Node**

$L_i$ State	$N_i$				$\bar{N}_i$			
	$I_i$		$\bar{I}_i$		$I_i$		$\bar{I}_i$	
	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$
$N_i$	0	0	0	1	0	0	0	0
$I_i$	0	0	0	0	0	1	0	0
$X_i$	0	0	0	0	0	0	1	0
$NA$	1	1	1	0	1	0	0	1

The Conditional Probability Table for the virtual evidence node,  $V_i$ , and is instantiated True, is:

**Table 54: CPT for Virtual Evidence Node**

$VE_i$ State	$N_i$	$I_i$	$X_i$	NA
<b>True</b>	$\frac{Q_{E,i}}{2}$	$\frac{1 - Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 - \eta_i Q_{E,i}}{2}$	0
<b>False</b>	$1 - \frac{Q_{E,i}}{2}$	$\frac{Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 + \eta_i Q_{E,i}}{2}$	1

The Conditional Probability Table for the local cause condition node,  $C_i$  is:

**Table 55: CPT for Local Cause Condition Node**

$C_i$ State	$N_i$	$I_i$	$X_i$	NA
$C_i$	0	1	1	0
$\bar{C}_i$	1	0	0	1

### 3.2. Calculate Control Node States

Prior to the updating with the virtual evidence, the probability that the control node is within each state is:

$$\begin{aligned}
 P(L_i = X_i) &= \sum_j \sum_k \sum_y P(L_i = X_i | N_{i,j}, I_{i,k}, X_y) P(N_{i,j}) P(I_{i,k}) P(X_y) \\
 &= \{1\} \{Q_{E,i}\} \{1 - (1 - \eta_i) Q_{E,i}\} \{\eta_i Q_{E,i}\} + \{0\} \dots + \{0\} \\
 &= \eta_i Q_{E,i}^2 (1 - (1 - \eta_i) Q_{E,i})
 \end{aligned}$$

$$\begin{aligned}
P(L_i = I_i) &= \sum_j \sum_k \sum_y P(L_i = X_i | N_{i,j}, I_{i,k}, X_y) P(N_{i,j}) P(I_{i,k}) P(X_y) \\
&= \{1\} \{Q_{E,i}\} \{(1 - \eta_i) Q_{E,i}\} \{1 - \eta_i Q_{E,i}\} + \{0\} \dots + \{0\} \\
&= Q_{E,i}^2 (1 - \eta_i) (1 - \eta_i Q_{E,i})
\end{aligned}$$

$$\begin{aligned}
P(L_i = N_i) &= \sum_j \sum_k \sum_y P(L_i = N_i | N_{i,j}, I_{i,k}, X_y) P(N_{i,j}) P(I_{i,k}) P(X_y) \\
&= \{1\} \{1 - Q_{E,i}\} \{1 - (1 - \eta_i) Q_{E,i}\} \{1 - \eta_i Q_{E,i}\} + \{0\} \dots + \{0\} \\
&= (1 - Q_{E,i}) (1 - (1 - \eta_i) Q_{E,i}) (1 - \eta_i Q_{E,i})
\end{aligned}$$

$$\begin{aligned}
P(L_i = NA) &= \sum_j \sum_k \sum_y P(L_i = NA | N_{i,j}, I_{i,k}, X_y) P(N_{i,j}) P(I_{i,k}) P(X_y) \\
&= 0
\end{aligned}$$

Each state of the control node can be updated given the virtual evidence using the following formula:

$$P(L_i | V_i) = \frac{P(L_i) P(V_i | L_i)}{\sum_j P(L_{i,j}) P(V_i | L_{i,j})}$$

The denominator is a normalization constant which can be calculated as  $\tau$ :

$$\begin{aligned}
\tau &= \sum_j P(L_{i,j})P(V_A|L_{i,j}) \\
&= \eta_i Q_{E,i}^2 (1 - (1 - \eta_i)Q_{E,i}) \left\{ \frac{1 - \eta_i Q_{E,i}}{2} \right\} \\
&\quad + Q_{E,i}^2 (1 - \eta_i)(1 - \eta_i Q_{E,i}) \left\{ \frac{Q_{E,i}(\eta_i - 1) + 1}{2} \right\} \\
&\quad + (1 - Q_{E,i})(1 - (1 - \eta_i)Q_{E,i})(1 - \eta_i Q_{E,i}) \left\{ \frac{Q_{E,i}}{2} \right\} \\
\tau &= \frac{Q_{E,i}(1 - \eta_i Q_{E,i})(1 - Q_{E,i}(1 - \eta_i))}{2}
\end{aligned}$$

The probability for each state of  $L_i$  given virtual evidence is:

$$\begin{aligned}
P(L_i = X_i|V_i) &= \frac{P(L_i = X_i)P(V_i|L_i = X_i)}{\tau} \\
&= \frac{2\eta_i Q_{E,i}^2 (1 - Q_{E,i}(1 - \eta_i)) (1 - \eta_i Q_{E,i})}{2Q_{E,i} (1 - Q_{E,i}(1 - \eta_i)) (1 - \eta_i Q_{E,i})} \\
&= \eta_i Q_{E,i} \\
P(L_i = I_i|V_i) &= \frac{P(L_i = I_i)P(V_i|L_i = I_i)}{\tau} \\
&= \frac{2Q_{E,i}^2 (1 - \eta_i)(1 - \eta_i Q_{E,i}) (1 - Q_{E,i}(1 - \eta_i))}{2Q_{E,i} (1 - Q_{E,i}(1 - \eta_i)) (1 - \eta_i Q_{E,i})} \\
&= Q_{E,i}(1 - \eta_i)
\end{aligned}$$



$$\begin{aligned}
P(L_i = N_i|V_i) &= \frac{P(L_i = N_i)P(V_i|L_i = N_i)}{\tau} \\
&= \frac{2(1 - Q_{E,i})(1 - (1 - \eta_i)Q_{E,i})(1 - \eta_i Q_{E,i})Q_{E,i}}{2Q_{E,i}(1 - Q_{E,i}(1 - \eta_i))(1 - \eta_i Q_{E,i})} \\
&= (1 - Q_{E,i})
\end{aligned}$$

$$\begin{aligned}
P(L_i = NA|V_i) &= \frac{P(L_i = NA)P(V_i|L_i = NA)}{\tau} \\
&= 0
\end{aligned}$$

### 3.3. Summary

This appendix has shown algebraically that the probability for each state of the control node is equivalent to the probability of the mutually exclusive parent nodes.

$$\begin{aligned}
P(L_i = X_i|V_i) &= \eta_i Q_{E,i} = P(X_i) \\
P(L_i = I_i|V_i) &= Q_{E,i}(1 - \eta_i) = P(I_i) \\
P(L_i = N_i|V_i) &= (1 - Q_{E,i}) = P(N_i) \\
P(L_i = NA|V_i) &= 0
\end{aligned}$$

## Appendices 4: Calculation of Event Assessment for GDM

### Example 1

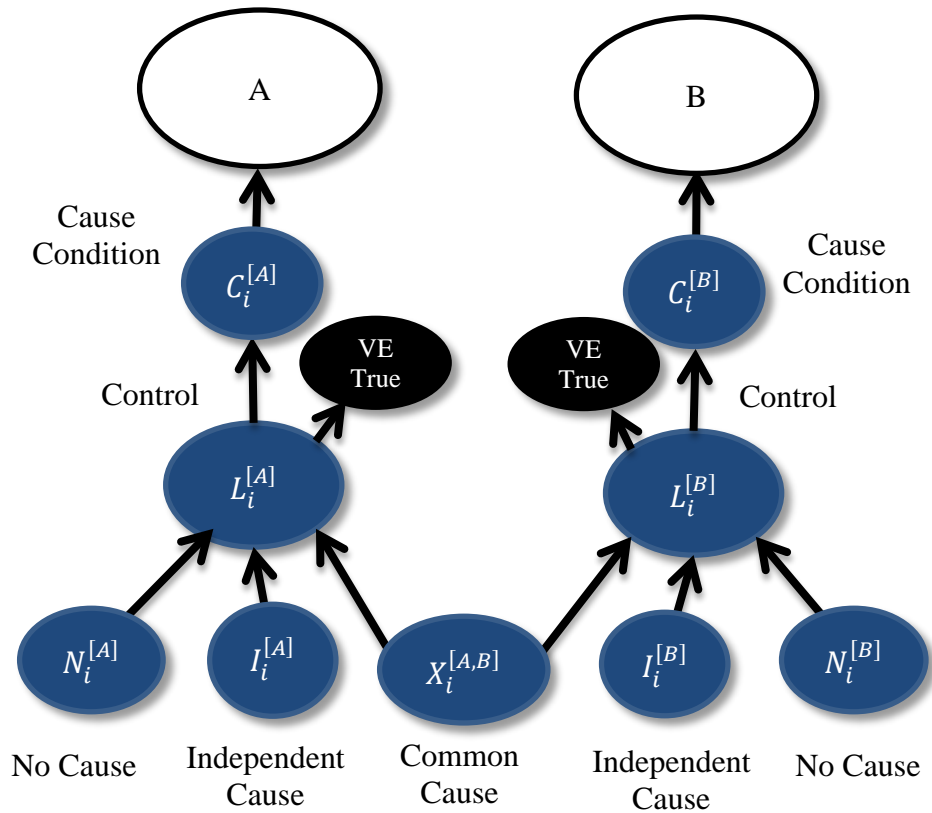
The following appendix shows algebraically the probability of a component B failing,

$P(B)$ , given knowledge that component A has failed from cause  $i$ ;  $P(B_i | C_i^{[A]})$ .

#### 4.1. The Bayesian Network

The Bayesian Network which provides modeling of the cause conditions is shown in

Figure 69.



**Figure 69: Structure of mutually exclusive Bayesian network with VE**

The Conditional Probability Tables for the three parent nodes,  $X_i, I_i, N_i$  are:

**Table 56: CPT for Common Cause Condition  $X_i$**

$X_i$ State	
$X_i$	$\eta_i Q_{E,i}$
$\bar{X}_i$	$1 - \eta_i Q_{E,i}$

**Table 57: CPT for Independent Cause Condition  $I_i$**

$I_i$ State	
$I_i$	$(1 - \eta_i) Q_{E,i}$
$\bar{I}_i$	$1 - (1 - \eta_i) Q_{E,i}$

**Table 58: CPT for No Cause Condition  $N_i$**

$N_i$ State	
$N_i$	$1 - Q_{E,i}$
$\bar{N}_i$	$Q_{E,i}$

The Conditional Probability Table for the control node,  $L_i$  is:

**Table 59: CPT for Control Node**

$L_i$ State	$N_i$				$\bar{N}_i$			
	$I_i$		$\bar{I}_i$		$I_i$		$\bar{I}_i$	
	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$	$C_i$	$\bar{C}_i$
$N_i$	0	0	0	1	0	0	0	0
$I_i$	0	0	0	0	0	1	0	0
$X_i$	0	0	0	0	0	0	1	0
$NA$	1	1	1	0	1	0	0	1

The Conditional Probability Table for the virtual evidence node,  $V_i$ , and is instantiated True, is:

**Table 60: CPT for Virtual Evidence Node**

$VE_i$ State	$N_i$	$I_i$	$X_i$	NA
True	$\frac{Q_{E,i}}{2}$	$\frac{1 - Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 - \eta_i Q_{E,i}}{2}$	0
False	$1 - \frac{Q_{E,i}}{2}$	$\frac{Q_{E,i}(1 - \eta_i)}{2}$	$\frac{1 + \eta_i Q_{E,i}}{2}$	1

The Conditional Probability Table for the local cause condition node,  $C_i$  is:

**Table 61: CPT for Local Cause Condition Node**

$C_i$ State	$N_i$	$I_i$	$X_i$	NA
$C_i$	0	1	1	0
$\bar{C}_i$	1	0	0	1

#### 4.2. Calculate Evidence Propagation

Table 62 to Table 70 show the marginal distribution values for each state of the nodes.

Note that the subscribe used to denote cause,  $i$ , has been omitted for brevity. All parameters are for a single cause.

**Table 62: Cause Condition Node ( $C^{[A]}$ )**

<b>State</b> ( $C_i^{[A]}$ )	$P(C_i^{[A]} V_1^{[A]})$	$P(C_i^{[A]} C_i^{[A]}, V_1^{[A]})$
<b>Equation</b>	$\sum_j P(C_i^{[A]} L_j^{[A]})P(L_j^{[A]} V_1^{[A]})$	$P(C_i^{[A]} C_i^{[A]}, V_1^{[A]})$
<b>Cause</b> ( $C_1^{[A]}$ )	$Q_E$	1
<b>No Cause</b> ( $C_2^{[A]}$ )	$1 - Q_E$	0

**Table 63: Virtual Evidence Node ( $V^{[A]}$ )**

<b>State</b> ( $V_i^{[A]}$ )	$P(V_i^{[A]} V_1^{[A]})$
<b>Equation</b>	
$V_1^{[A]}$	1
$V_2^{[A]}$	0

**Table 64: Control Node ( $L^{[A]}$ ):**

State ( $L_j^{[A]}$ )	$P(L_j^{[A]})$	$P(L_j^{[A]} V_1^{[A]})$	$P(L_j^{[A]} V_1^{[A]}, C_1^{[A]})$
<b>Equation</b>	$\sum_k \sum_y \sum_x \{P(N_k^{[A]})P(I_y^{[A]})P(X_x) * P(L_j^{[A]} N_k^{[A]}, I_y^{[A]}, X_x)\}$	$\frac{P(L_j^{[A]})P(V_1^{[A]} L_j^{[A]})}{\sum_j P(L_j^{[A]})P(V_1^{[A]} L_j^{[A]})}$	$\frac{P(V_1^{[A]} L_j^{[A]})P(C_1^{[A]} L_j^{[A]})}{\sum_j P(V_1^{[A]} L_j^{[A]})P(C_1^{[A]} L_j^{[A]})}$
<b>Nil</b> ( $L_1^{[A]}$ )	$(\eta Q_E - 1)(Q_E - 1)(Q_E(\eta - 1) + 1)$	$1 - Q_E$	0
<b>Ind</b> ( $L_2^{[A]}$ )	$Q_E^2(Q_E\eta - 1)(\eta - 1)$	$(1 - \eta)Q_E$	$1 - \eta$
<b>CC</b> ( $L_3^{[A]}$ )	$\eta Q_E^2(Q_E(\eta - 1) + 1)$	$\eta Q_E$	$\eta$
<b>NA</b> ( $L_4^{[A]}$ )	0	0	0

**Table 65: No Cause Condition Node ( $N^{[A]}$ )**

State ( $N_k^{[A]}$ )	$P(N_k^{[A]})$	$P(N_k^{[A]} C_1^{[A]})$
<b>Equation</b>	$P(N_k^{[A]})$	$\frac{P(N_k^{[A]}) \sum_j P(C_1^{[A]} L_j^{[A]}) P(L_j^{[A]} N_k^{[A]})}{\sum_k P(N_k^{[A]}) \sum_j P(C_1^{[A]} L_j^{[A]}) P(L_j^{[A]} N_k^{[A]})}$
<b>Nil Cause</b> ( $N_1^{[A]}$ )	$1 - Q_E$	0
<b>Cause</b> ( $N_2^{[A]}$ )	$Q_E$	1

**Table 66: Independent Cause Condition Node ( $I^{[A]}$ )**

State ( $I_y^{[A]}$ )	$P(I_y^{[A]})$	$P(I_y^{[A]}   C_1^{[A]})$
<b>Equation</b>	$P(I_y^{[A]})$	$\frac{P(I_y^{[A]}) \sum_j P(C_1^{[A]}   L_j^{[A]}) P(L_j^{[A]}   I_y^{[A]})}{\sum_y P(I_y^{[A]}) \sum_j P(C_1^{[A]}   L_j^{[A]}) P(L_j^{[A]}   I_y^{[A]})}$
$I_1^{[A]}$	$(1 - \eta)Q_E$	$1 - \eta$
$I_2^{[A]}$	$1 - (1 - \eta)Q_E$	$\eta$

**Table 67: Common Cause Condition Node ( $X_x$ )**

State ( $X_x$ )	$P(X_x)$	$P(X_x   C_1^{[A]})$
<b>Equation</b>	$P(X_x)$	$\frac{P(X_x) \sum_j P(C_1^{[A]}   L_j^{[A]}) P(L_j^{[A]}   X_x)}{\sum_x P(X_x) \sum_j P(C_1^{[A]}   L_j^{[A]}) P(L_j^{[A]}   X_x)}$
$X_1$	$\eta Q_E$	$\eta$
$X_2$	$1 - \eta Q_E$	$1 - \eta$

**Table 68: Control Node ( $L^{[B]}$ )**

State ( $L_j^{[B]}$ )	$P(L_j^{[B]})$	$P(L_j^{[B]} V_1^{[B]})$	$P(L_j^{[B]} V_1^{[B]}, C_1^{[A]})$
<b>Equation</b>	$\sum_k \sum_y \sum_x \{P(N_k^{[B]})P(I_y^{[B]})P(X_x) * P(L_j^{[B]} N_k^{[B]}, I_y^{[B]}, X_x)\}$	$\frac{P(L_j^{[B]})P(V_1^{[B]} L_j^{[B]})}{\sum_j P(L_j^{[B]})P(V_1^{[B]} L_j^{[B]})}$	$\frac{P(L_j^{[B]} C_1^{[A]})P(V_1^{[B]} L_j^{[B]})}{\sum_j P(L_j^{[B]} C_1^{[A]})P(V_1^{[B]} L_j^{[B]})}$
<b>Nil</b> ( $L_1^{[B]}$ )	$(\eta Q_E - 1)(Q_E - 1)(Q_E(\eta - 1) + 1)$	$1 - Q_E$	$(1 - \eta) \cdot \frac{Q_E - 1}{\eta Q_E - 1}$
<b>Ind</b> ( $L_2^{[B]}$ )	$Q_E^2(Q_E\eta - 1)(\eta - 1)$	$(1 - \eta)Q_E$	$\frac{Q_E(\eta - 1)^2}{1 - Q_E - Q_E(\eta - 1)}$
<b>CC</b> ( $L_3^{[B]}$ )	$\eta Q_E(Q_E(\eta - 1) + 1)$	$\eta$	$\eta$
<b>NA</b> ( $L_4^{[B]}$ )	0		0



**Table 69: Cause Condition Node ( $C^{[B]}$ )**

State ( $C_i^{[B]}$ )	$P(C_i^{[B]} V_1^{[B]})$	$P(C_1^{[B]} V_1^{[B]}, C_1^{[A]})$
<b>Equation</b>	$\sum_j P(C_i^{[B]} L_j^{[B]}) P(L_j^{[B]} V_1^{[B]})$	$\sum_{L_A=L_A} P(C_1^{[B]} L_j^{[B]}) P(L_j^{[B]} V_1^{[B]}, C_1^{[A]})$
<b>Cause (<math>C_1^{[B]}</math>)</b>	$Q_E$	$\eta + \frac{Q_E(\eta - 1)^2}{1 - Q_E - Q_E(\eta - 1)}$
<b>No Cause (<math>C_2^{[B]}</math>)</b>	$1 - Q_E$	$1 - \eta - \frac{Q_E(\eta - 1)^2}{1 - Q_E - Q_E(\eta - 1)}$

**Table 70: Second Component Node ( $B$ )**

State ( $B$ )	$P(B V_1^{[B]})$	$P(B V_1^{[B]}, C_1^{[A]})$
<b>Equation</b>	$\sum_i P(B C_i^{[B]}) P(C_i^{[B]} V_1^{[B]})$	$\sum_i P(B C_i^{[B]}) P(C_i^{[B]} V_1^{[B]}, C_1^{[A]})$
<b><math>B</math></b>	$p_i Q_E$	$p_i \eta + \frac{p_i Q_E(\eta - 1)^2}{1 - Q_E - Q_E(\eta - 1)}$
<b><math>\bar{B}</math></b>	$1 - p_i Q_E$	$1 - p_i \eta - \frac{p_i Q_E(\eta - 1)^2}{1 - Q_E - Q_E(\eta - 1)}$

### 4.3. Summary

This appendix has shown algebraically the probability of a component B failing,  $P(B)$ ,

given knowledge that component A has failed from cause  $i$ ;  $P(B_i|C_i^{[A]})$ .

$$P(B_i|C_i^{[A]}) = p_i\eta_i + \frac{p_iQ_{E,i}(\eta_i - 1)^2}{1 - Q_{E,i} - Q_{E,i}(\eta_i - 1)}$$

## Notation

### General notation

$P(X)$	The probability of event $X$
$X^{[E]}$	A parameter which is related to component type 'E'.
$A_i$	The independent failure of event $A$
$X_{AB}$	The common cause failure event of components $A$ and $B$ .
$m$	The size of a common cause component group.
$k$	A multiplicity of failure within a common cause component group. ( $k$ failures out of $m$ components)
$w$	The number of coupling factor features which are being assessed in the target system.

### Component event count parameters

$Q_T$	The total failure frequency/probability of each component due to independent and common cause events.
$\beta$	A parameter of the Beta Factor Model. This is the portion of an individual component's failure probability/rate which is a common cause failure.
$N_1$	The total number of component demands. Assuming that each time the system is demanded, all components are demanded the following relationship exists between $N_D$ and $N_1$ : $N_k = mN_D$
$n_F$	The total number of component failures. $n_F = \sum_{k=1}^m kn_k$
$p_i$	Component fragility to cause $i$ . (GDM parameter)

$\eta_i$	Coupling factor strength for cause $i$ . (GDM parameter)
$Q_{E,i}$	Cause condition probability for cause $i$ . (GDM parameter)
$Q_{IE,i}$	The independent cause condition probability for a component $Q_{IE,i} = (1 - \eta_i)Q_{E,i}$
$Q_{CE,i}$	The common cause condition probability for a component. $Q_{CE,i} = \eta_i Q_{E,i}$

CCF event count parameters

$I_h$	the $h^{th}$ hypotheses for an observed CCF event. Where $1 \leq h \leq H$ .
$\bar{I}$	the average impact vector for a CCF event. This is the weighted sum of all hypotheses for the event. $\bar{I} = \sum_{h=1}^H w_h I_h$
$\bar{F}_k(j)$	the $k^{th}$ element of the average impact vector where $(0 \leq k \leq m)$ for the $j^{th}$ event where $(0 \leq j \leq J)$
$n_k$	the total number of CCF basic events involving the failure of $k$ components within a CCCG. $n_k = \sum_{j=1}^J F_k(j)$
$n_t$	the total number of common cause failure events. $n_t = \sum_{k=1}^m n_k$
$N_D$	The number of demands on the CCCG (assuming that each time the system is demanded, all components are demanded) $N_D = \binom{m}{k}^{-1} N_k = n_0 + n_t$
$N_k$	The number of demands on a subset group of components within the CCCG of size $k$ . (assuming that each time the system is demanded, all components are demanded)

$$N_k = \binom{m}{k} N_D$$

$Q_k^{(m)}$  Basic event failure frequency/probability for  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ). This is a parameter of the Basic Parameter Model. For example, in a three train system of components A, B and C:

$$\begin{aligned} Q_1^{(3)} &= P(A_i) = P(B_i) = P(C_i) \\ Q_2^{(3)} &= P(X_{AB}) = P(X_{AC}) = P(X_{BC}) \\ Q_3^{(3)} &= P(X_{ABC}) \end{aligned}$$

The basic parameter estimator is:

$$Q_k^{(m)} = \frac{n_k}{N_k}$$

$\alpha_k$  An alpha factor which is a parameter of the alpha factor model. This is the fraction of CCF failure event where  $k$  components fail within the CCCG. ( $1 \leq k \leq m$ ).

$n_{k,i}$  the number of failure events/frequency which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ) of coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$ .

$n_{p,i}$  the total number of failure events/frequency which had the opportunity for the failure to propagate through coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$ .

$n_{t,i}$  the total number of common cause failure events for coupling factor/cause  $i$  where  $i \in \{1,2,3,\dots,w\}$ .

$\alpha_{k,i}$  a partial alpha factor which represents the portion of system failure events which resulted in  $k$  components failing within a common cause component group of size  $m$ , ( $1 \leq k \leq m$ ) when there was a potential for failure propagation through coupling factor  $i$  where  $i \in \{1,2,3,\dots,w\}$ .

$a, b$  Beta distribution parameters for the distribution of parameter  $\theta$  where  $\theta \sim \text{Beta}(a, b)$ . This is used for Bayesian estimators of parameters which range from  $0 \leq \theta \leq 1$ .

$\alpha'_k$  the assessed alpha factor. This is the system alpha factor which only considers the coupling factors shared by the components within the CCCG.

- $\gamma$  the portion of failure events for each cause  $[\gamma_1, \gamma_2, \dots, \gamma_w]$
- $\varphi$  the equivalent count of failure events for each cause  $[\varphi_1, \varphi_2, \dots, \varphi_w]$

## Glossary

<i>Basic Event</i>	An event in a reliability logic model that represents the state in which a component or group of components is unavailable and does not require further development in terms of contributing cases.
<i>Common Cause Event</i>	An unexpected dependent failure in which two or more component fault states exist simultaneously, or within a short time interval, and are a direct result of a shared cause. The failure is classified as unexpected if it has not been explicitly modeled within a PRA.
<i>Common Cause Basic Event</i>	In system modeling, a basic event that represents the unavailability of a specific set of components because of shared causes that are not explicitly represented in the system logic model as other basic events.
<i>Common Cause Component Group</i>	A group (usually similar in mission, manufacturer, maintenance, environment, etc) of components that are considered to have a high potential for failure due to the same cause or causes.
<i>Common Cause Failure Model</i>	The basis for quantifying the frequency of common cause events. Examples include the beta factor, alpha factor, basic parameter, and the binomial failure rate models.
<i>Component</i>	An element of plant hardware designed to provide a particular function.
<i>Component Boundary</i>	The component boundary encompasses the set of piece parts that are considered to form the component.
<i>Component State</i>	Component state defines the component status in regard to its intended function. Two general categories of component states are <i>available</i> and <i>unavailable</i> .
<i>Conditional Probability Table</i>	The CPT defined the probability of each state of a Bayesian Network node, conditional on the state of parent nodes.
<i>Coupling Factor/Mechanism</i>	A system feature which is shared by multiple components such that it creates a dependency.

<i>Defense</i>	Any operational, maintenance, and design measures taken to diminish the frequency and/or consequences of common cause failures.
<i>Event</i>	An event is the occurrence of a component state or group of component states.
<i>Event (CCF Perspective)</i>	The count process as observed at the CCCG level. Multiple failures from a single demand are considered one event.
<i>Event (Component Perspective)</i>	The count process as observed at the component level. A single CCF event from a CCF Event Perspective will have multiple demands and failures from a Component Level Perspective.
<i>Failure Mechanism</i>	The history describing the event and influences leading to a given failure.
<i>Failure Mode</i>	A description of the component failure in terms of the component function that was actually or potentially unavailable.
<i>Impact Vector</i>	As assessment of the impact an event (CCF Event Perspective) would have on a common cause component group. The impact is usually measured as the number of failed components out of a set of similar components in the common cause component group.
<i>Independent Basic Events</i>	Two basic events, A and B, are statistically independent if, $P(B A) = P(B)$ therefore $P(A \cap B) = P(A)P(B)$ , where $P(X)$ denotes the probability of event X.
<i>Mutually Exclusive</i>	Events which cannot occur at the same time. A and B are mutually exclusive if $P(A \cap B) = 0$ .
<i>Rare Event Approximation</i>	Where the following approximation is used, $P(A \cup B) \cong P(A) + P(B)$ instead of the correct formula $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ . This can be used where $P(A \cap B) \ll P(A)$ and $P(A \cap B) \ll P(B)$ .
<i>Soft Dependencies</i>	Soft dependencies have a probabilistic relationship. For example if one component failed, there is a probability that



a second component could fail through a shared cause such as a manufacturer error, or shared maintenance procedure

*Soft Evidence*

Uncertain evidence used in the ‘all things considered’ method. Soft Evidence is treated with Jeffrey’s Rule. (Darwiche 2009)

*Train*

A train is a success path for a system. A two train system has two paths in redundancy.

*Uncertain Evidence*

Evidence where the observer has a distribution of beliefs over the possible outcomes. This includes virtual evidence and soft evidence.

Virtual Evidence

Uncertain evidence used in the ‘nothing else considered’ method. (Darwiche 2009)

## Abbreviations

AFM	Alpha Factor Model
BFR / BFRM	Binomial Failure Rate Model
BN	Bayesian Network
BP	Basic Parameter (a parameter to the Basic Parameter Model)
CCBE	Common Cause Basic Event
CCCG	Common Cause Component Group
CCF	Common Cause Failure
CCFDB	Common Cause Failure Database (NRC)
CPT	Conditional Probability Table
EDG	Emergency Diesel Generator
RBD	Reliability Block Diagram
PRA	Probability Risk Assessment
UOI	Unknown Of Interest

## Bibliography

- Anude, 1994. *The analysis of redundant reliability systems with common-cause failures*. Canada: University of Ottawa.
- Apostolakis, G.E. & Moieni, P., 1987. The Foundation of Models of Dependence in Probability Safety Assessments. *Reliability Engineering*, 18, p.177–195.
- Atwood, C.L., 1996. Constrained noninformative priors in risk assessment. *Reliability Engineering & System Safety*, 53(1), p.37–46.
- Atwood, C.L., 1986. The binomial failure rate common cause model. *Technometrics*, 28(2), p.139–148.
- Bates, A., 1995. *Use of Probabilistic Risk Assessment Methods in Nuclear Regulatory Activities; Final Policy Statement*, US Nuclear Regulatory Commission.
- Bell, J. & Holroyd, J., 2009. *Review of human reliability assessment methods*, Health and Safety Executive.
- Bourne, A.J., Edwards, G.T. & Hunns, D.M., 1981. *Defences against common-mode failures in redundancy systems: a guide for management designers and operators*, UKAEA Safety and Reliability Directorate.
- Brand, P.V. & Gabbot, D., 1993. Unified Partial Method for dependent failures assessment. *Technical Report AEA Technology*.
- Cooke, R.M., 1991. *Experts in uncertainty*, Oxford University Press US.
- Darwiche, A., 2009. *Modeling and Reasoning with Bayesian Networks* 1st ed., Cambridge University Press.
- Dörre, P., 1989. Basic aspects of stochastic reliability analysis for redundancy systems. *Reliability Engineering & System Safety*, 24(4), p.351–375.
- Edwards, G.T. & Watson, I.A., 1979. *A study of common-mode failures*, U.K.A.E.A.
- Ericson II, C.A., 2005. *Hazard Analysis Techniques for System Safety* 1st ed., Wiley-Interscience.
- Evans, M.G.K., Parry, G.W. & Wreathall, J., 1984. On the treatment of common-cause failures in system analysis. *Reliability engineering*, 9(2), p.107–115.

- Fenton, N., Neil, M. & Lagnado, D., 2012. Modelling mutually exclusive causes in Bayesian networks. *Draft*.
- Fleming, K.N., 1975. A Reliability Model for Common Mode Failure in Redundant Safety Systems. In *Proceedings of the Sixth Annual Pittsburgh Conference on Modeling and Simulation*. Pittsburgh.
- Fleming, K.N. & Kalinowski, A.M., 1983. *An extension of the beta factor method to systems with high levels of redundancy* 6th ed., Pickard, Lowe and Garrick Inc.
- Fleming, K.N. & Mosleh, A., 1985. *Classification and analysis of reactor operating experience involving dependent events*, Pickard, Lowe and Garrick, Inc., Newport Beach, CA (USA).
- Fleming, K.N., Mosleh, A. & Kelley Jr, A.P., 1983. Analysis of dependent failures in risk assessment and reliability evaluation. *Nucl. Saf.:(United States)*, 24(5).
- Garthwaite, P.H., Kadane, J.B. & O'Hagan, A., 2005. Statistical methods for eliciting probability distributions. *Journal of the American Statistical Association*, 100(470), p.680–701.
- Guey, C.N., 1984. A method for estimating common cause failure probability and model parameters : the inverse stress-strength interference (ISSI) technique.
- Hakansson, M., 2011. *Feasibility study of a strength of defence method for estimation of CCF probabilities*, Uppsala University.
- Han, S.G., Yoon, W.H. & Chang, S.H., 1989. The trinomial failure rate model for treating common mode failures. *Reliability Engineering & System Safety*, 25(2), p.131–146.
- Harris, B., 1986. Stochastic Models for Common Cause Failures. In *Proceedings of the International Conference on Reliability and Quality Control*. Conference on Reliability and Quality Control. pp. 185–200.
- Hauptmanns, U., 1996. The multi-class binomial failure rate model. *Reliability Engineering & System Safety*, 53(1), p.85–90.
- Hirschberg, S., 1985. Comparison of Methods for Quantitative Analysis of Common Cause Failures- A Case Study. In *Proceedings of ANS/ENS International Tropical Meeting on Probabilistic Safety Methods and Applications*. San Francisco, California: Electric Power Research Institute, pp. 183/1–183/10.
- Hokstad, P., 2004. A Generalisation of the Beta Factor Model, Probabilistic Safety Assessment and Management. In *Proceedings from PSAM7-ESREL*. PSAM7.

Springer.

- Hokstad, P., 1988. A shock model for common-cause failures. *Reliability Engineering & System Safety*, 23(2), p.127–145.
- Hokstad, P., Maria, A. & Tomis, P., 2006. Estimation of Common Cause Factors From Systems With Different Numbers of Channels. *IEEE Transactions on Reliability*, 55(1), p.18–25.
- Hokstad, P. & Rausand, M., 2008. Common cause failure modeling: Status and trends. *Handbook of Performability Engineering*, p.621–640.
- Hughes, R.P., 1987. A new approach to common cause failure. *Reliability Engineering*, 17(3), p.211–236.
- Jo, Y., 2005. Modeling and quantification of common cause failures among pumps with different operation histories. In *Proceedings (CD) of Topical Meeting PSA '05*. PSA. San Francisco, California: American Nuclear Society, pp. 1375–1382.
- Johnston, B.D., 1987. A structured procedure for dependent failure analysis (DFA). *Reliability Engineering*, 19(2), p.125–136.
- Kadane, J. & Wolfson, L.J., 1998. Experiences in elicitation. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1), p.3–19.
- Kang, D.I. et al., 2009. Approximate formulas for treating asymmetrical common cause failure events. *Nuclear Engineering and Design*, 239(2), p.346–352.
- Kass, R.E. & Wasserman, L., 1996. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435), p.1343–1370.
- Kaufman, L.M., Bhide, S. & Johnson, B.W., 2000. Modeling of common-mode failures in digital embedded systems. In *Reliability and Maintainability Symposium, 2000. Proceedings. Annual*. pp. 350–357.
- Kelly, D. et al., 2011. Common-Cause Failure Treatment in Event Assessment: Basis for a Proposed Model. In *DRAFT*.
- Korb, K.B. & Nicholson, A.E., 2004. *Bayesian artificial intelligence*, CRC Press.
- Kreuser, A. & Peschke, J., 1997. Coupling model: A common-cause-failure model with consideration of interpretation and projection uncertainties. In *Proceedings of Jahrestagung Kerntechnik*. Jahrestagung Kerntechnik. Aachen, Germany.

- Kreuser, A. & Peschke, J., 2001. Coupling model: A common-cause-failure model with consideration of interpretation uncertainties. *Nuclear technology*, 136(3).
- Kvam, P.H., 1998a. A parametric mixture-model for common-cause failure data [of nuclear power plants]. *Reliability, IEEE Transactions on*, 47(1), p.30–34.
- Kvam, P.H., 1993. *Computational problems with the binomial failure rate model and incomplete common cause failure reliability data*, Los Alamos National Lab., NM (United States).
- Kvam, P.H., 1998b. The binomial failure rate mixture model for common cause failure data from the nuclear industry. *Journal of the Royal Statistical Society*, 47(1), p.49–61.
- Kvam, P.H. & Martz, H.F., 1995. Bayesian inference in a discrete shock model using confounded common cause data. *Reliability Engineering & System Safety*, 48(1), p.19–25.
- Lindberg, S., 2007. *Common Cause Failure Analysis: Methodology evaluation using Nordic experience data*. Uppsala, Sweden: Uppsala Universitet.
- Mankamo, T., 1977. *Common load model: a tool for common cause failure analysis*, Valtion Teknillinen Tutkimuskeskus, Espoo (Finland). Saehkoetekniikan Lab.
- Marshall, A.W. & Olkin, I., 1967. A multivariate exponential distribution. *Journal of the American Statistical Association*, 62(317), p.30–44.
- Marshall, F.M., Rasmuson, D.M. & Mosleh, A., 1998. *Common-Cause Failure Parameter Estimations*, U.S. Nuclear Regulatory Commission.
- Martin, B.R. & Wright, R., 1987. A Practical Method of Common Cause Failure Modeling. *Reliability Engineering*, 19, p.185–199.
- Meyer, M.A. & Booker, J.M., 1990. *Eliciting and Analyzing Expert Judgement, A Practical Guide*, Washington DC: US Nuclear Regulatory Commission.
- Mohagheh, Z., Kazemi, R. & Mosleh, A., 2009. Incorporating organizational factors into Probabilistic Risk Assessment (PRA) of complex socio-technical systems: A hybrid technique formalization. *Reliability Engineering & System Safety*, 94(5), p.1000–1018.
- Mohagheh, Z., Modarres, M. & Christou, A., 2011. Physics-Based Common Cause Failure Modeling in Probabilistic Risk Analysis: A Mechanistic Perspective. In *Proceedings of the ASME 2011 Power Conference*. POWER2011. Denver, Colorado, USA.

- Mohaghegh, Z. & Mosleh, A., 2009. Incorporating organizational factors into probabilistic risk assessment of complex socio-technical systems: Principles and theoretical foundations. *Safety Science*, 47(8), p.1139–1158.
- Mosleh, A., 1991. Common cause failures: an analysis methodology and examples. *Reliability Engineering & System Safety*, 34(3), p.249–292.
- Mosleh, A., 1986. Hidden sources of uncertainty: Judgment in the collection and analysis of data. *Nuclear Engineering and Design*, 93(2–3), p.187–198.
- Mosleh, A. et al., 1988. *Procedures for Treating Common Cause failures in Safety and Reliability Studies*, U.S. Nuclear Regulatory Commission.
- Mosleh, A. & Goldfeiz, E., 1997. *An Approach for Assessing The Impact of Organizational Factors on Risk*, US Nuclear Regulatory Commission.
- Mosleh, A., Rasmuson, D.. & Marshall, F.M., 1998. *Guidelines on Modeling Common Cause Failures in Probabilistic Risk Assessments*, Washington DC: U.S. Nuclear Regulatory Commission.
- Mosleh, A. & Siu, N.O., 1987. A Multi-Parameter, Event-Based Common Cause Failure Model. *Trans. 9th Int. Conf. Structural Mechanics in Reactor Technology* Lausanne, Switzerland, M, p.147.
- Mumpower, J.L. & Stewart, T.R., 1996. Expert judgement and expert disagreement. *Thinking & Reasoning*, 2(2-3), p.191–212.
- NASA, 2002. *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners*, NASA.
- NEA, 2011. *International Common Cause Failure Data Exchange (ICDE) General Coding Guidelines Updated Version*, Nuclear Energy Agency.
- Neapolitan, R.E., 2003. *Learning Bayesian Networks* illustrated edition., Prentice Hall.
- O'Hagan, A. et al., 2006. *Uncertain judgements: eliciting experts' probabilities*, Wiley Chichester.
- Parry, G.W., 1989. Comments On: Basic aspects of stochastic reliability analysis for redundancy systems. *Reliability Engineering & System Safety*, 24(4), p.377–381.
- Paula, H., 1995. Technical Note: On the definition of common-cause failures. *Nuclear Safety*, 36(1).

- Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann San Mateo, CA.
- Rasmuson, D.M., Burdick, G.R. & Wilson, J.H., 1979. *Common cause failure analysis techniques: a review and comparative evaluation*, Department of Energy, Idaho Operations Office, Idaho National Engineering Laboratory.
- Rasmussen, N., 1975. *Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*, Washington DC: U.S. Nuclear Regulatory Commission.
- Rausand, M. & Høyland, A., 2003. *System Reliability Theory: Models, Statistical Methods, and Applications, Second Edition* 2nd ed., Wiley-Interscience.
- Ronald, B. et al., 2005. *Simplified Expert Elicitation Guideline for Risk Assessment of Operating Events*, Idaho National Laboratory (INL).
- Siu, N.O. & Kelly, D.L., 1998. Bayesian parameter estimation in probabilistic risk assessment. *Reliability Engineering & System Safety*, 62(1–2), p.89–116.
- Skjong, R. & Wentworth, B.H., 2001. Expert judgment and risk perception. In *International Offshore and Polar Engineering Conference*. pp. 537–544.
- Smith, A.M. & Watson, I.A., 1980. Common cause failures—a dilemma in perspective. *Reliability Engineering*, 1(2), p.127–142.
- Stamatelatos, M.G., 1982. Improved Method for Evaluating Common-Cause Failure Probabilities. *Transactions of the American Nuclear Society on Probabilistic Risk Assessment*, 43, p.474–481.
- Steppeler, J. et al., 2003. Review of numerical methods for nonhydrostatic weather prediction models. *Meteorology and Atmospheric Physics*, 82(1-4), p.287–301.
- Vaurio, J.K., 1998. An implicit method for incorporating common-cause failures in system analysis. *Reliability, IEEE Transactions on*, 47(2), p.173–180.
- Vaurio, J.K., 2008. Common Cause Failure Modeling. *Encyclopedia of Quantitative Risk Analysis and Assessment*.
- Vaurio, J.K., 1999. Common-cause failure models, data, quantification. *Reliability, IEEE Transactions on*, 48(3), p.213–214.
- Vaurio, J.K., 2007. Consistent mapping of common cause failure rates and alpha factors. *Reliability Engineering & System Safety*, 92(5), p.628–645.



- Vaurio, J.K., 1981. Structures for Common Cause Failure Analysis. In PRA Meeting. Portchester, New York, pp. 676–685.
- Vesely, W.E., 1977. Estimating common cause failure probabilities in reliability and risk analysis: Marshall-Olkinspecializations. In: Fussell and Burdick Nuclear systems reliability engineering and risk assessment. In *Nuclear Systems Reliability Engineering and Risk*. SIAM. Philadelphia, pp. 314–341.
- Vesely, W.E., Uryasev, S.P. & Samanta, P.K., 1994. Failure of emergency diesel generators: a population analysis using empirical Bayes methods. *Reliability Engineering & System Safety*, 46(3), p.221–229.
- Werner, W., 1994. *Results of recent risk studies in France, Germany, Japan, Sweden and the United States*, Paris: OECD Nuclear Energy Agency.
- Wierman, T., 2013. Causes for Single Failures.
- Wierman, T. & Kvarfordt, K.J., 2011. Data Extraction from CCF Database.
- Wierman, T., Rasmuson, D. & Stockton, N., 2003a. *Common-Cause Failure Event Insights: Circuit Breakers*, Washington DC: U.S. Nuclear Regulatory Commission.
- Wierman, T., Rasmuson, D. & Stockton, N., 2003b. *Common-Cause Failure Event Insights: Emergency Diesel Generators*, Washington DC: U.S. Nuclear Regulatory Commission.
- Wierman, T., Rasmuson, D. & Stockton, N., 2003c. *Common-Cause Failure Event Insights: Motor-Operated Valves*, Washington DC: U.S. Nuclear Regulatory Commission.
- Wierman, T., Rasmuson, D. & Stockton, N., 2003d. *Common-Cause Failure Event Insights: Pumps*, Washington DC: U.S. Nuclear Regulatory Commission.
- Wierman, T., Rasmuson, D.M. & Mosleh, A., 2007. *Common-Cause Failure Database and Analysis System: Event Data Collection, Classification, and Coding*, U.S. Nuclear Regulatory Commission.
- Xie, L., 1998. A knowledge-based multi-dimension discrete common cause failure model. *Nuclear engineering and design*, 183(1), p.107–116.
- Yamaguchi, M. & Donn, J., 2011. In Japan plant, frantic efforts to avoid meltdown - Yahoo! News.
- Yang & Berger, 1998. A Catalog of Noninformative Priors (DRAFT).

Zitrou, A., 2006a. *Exploring a bayesian approach for structural modelling of common cause failures*. Department of Management Science: University of Strathclyde.

Zitrou, A., 2006b. *Exploring a Bayesian Approach for Structural Modelling of Common Cause Failures*. Dissertation. University of Strathclyde.