

Graph-based fault detection for a gas-to-liquids process: an exergy approach

S. Greyling

 orcid.org/0000-0002-2163-3611

Thesis submitted in fulfilment of the requirements for the degree *Doctor of Philosophy in Electrical and Electronic Engineering* at the North-West University

Promoter: Prof G. van Schoor

Co-promoter: Prof K.R. Uren

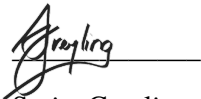
Co-promoter: Dr H. Marais

Examination: December 2020

Student number: 21818347

DECLARATION

I, Sarita Greyling, hereby declare that the thesis entitled “Graph-based fault detection for a gas-to-liquids process: an exergy approach” is my own original work and has not already been submitted to any other university or institution for examination.



Sarita Greyling
21818347

Signed on the 17th day of December 2020 at Potchefstroom

ACKNOWLEDGEMENTS

"Trust in the Lord with all your heart; do not depend on your own understanding. Seek His will in all you do, and He will show you which path to take. Do not be impressed with your own wisdom. Instead, fear the Lord and turn away from evil. Then you will have healing for your body and strength for your bones. Honour the Lord with your wealth and with the best part of everything you produce. Then He will fill your barns with grain, and your vats will overflow with good wine."

I want to wholeheartedly thank the following, in no specific order:

- The North-West University Potchefstroom Campus for providing me with the opportunity and financial support to enrol for a Doctor of Philosophy degree.
- My supervisor, Prof George van Schoor, for his unwavering support, insights, leadership, and inspiration.
- My co-supervisor, Prof Kenny Uren, for his encouragement, guidance, kindness, and patience.
- My second co-supervisor, Dr Henri Marais, for his constructive inputs, useful advice, and willing assistance.
- McTronX, our research group, for everyone's unique contribution and counsel.
- My husband, Werner Greyling, for all of his motivation, love and support throughout this study.
- My father, mother, and sister, for their ever-lasting nurturance.

ABSTRACT

As there are many safety and financial risks within modern process plants, process monitoring is said to be indispensable. Process monitoring aids operators in ensuring reliable and efficient operation of the plant. Fault detection and Isolation (FDI), which make up a large portion of a process monitoring protocol, is a sophisticated scheme which aims to detect and isolate anomalies that occur within the plant. For the past 50+ years, much work has been done on developing FDI schemes for a vast array of different applications. In recent years, novel energy-based FDI techniques were proposed, as energy is seen as a unifying parameter of different domains. These energy-based approaches also endeavour to capture causal (or structural) information of the physical system.

Keeping with this theme, this study will determine, after some alterations, the applicability and performance of some of the previously developed energy-based approaches, especially compared to one another, when applied to a single, larger-scale petrochemical process. The petrochemical process, a gas-to-liquids (GTL) process, is not seen within the FDI literature, and could arguably be well-suited to being used as a benchmark to evaluate the performance of proposed FDI schemes. A such, this study systematically documents the process specifics and modelling effort to allow easy recreation thereof. As the model was simulated within the commercial process simulator, Aspen HYSYS®, and the FDI approaches require energy data, *user variables* were created to compute the desired energy data automatically. The different techniques investigated were a fixed-threshold approach, a graph matching approach using a distance parameter, and graph matching approach utilising eigendecomposition. The approaches and their methodologies are shown and applied accordingly; using the same normal and faulty energy data of the GTL model. The results are then interpreted in terms of the approaches' ability to detect and isolate the pre-defined faults. Finally, the performance of the approaches is compared to determine the best performing technique.

This study shows the degrees of applicability of the examined energy-based FDI approaches. It also found the graph-based approach utilising the distance parameter showing the most promise, as fault locations could be distinguished. This, therefore, confirms not only the usefulness of expressing the system in terms of energy but that structural information is also retained.

Keywords: *Energy, Exergy, Fault detection and isolation, Gas-to-liquids, Graph-based*

TABLE OF CONTENTS

Declaration	i
Acknowledgements	ii
Abstract	iii
List of Figures	x
List of Tables	xii
Abbreviations	xvi
Nomenclature	xviii
1 Introduction	1
1.1 Motivation	1
1.2 Problem statement	3
1.3 Research aims and objectives	3
1.4 Research methodology	3
1.5 Contribution of research	4
1.6 Publications from the research	6
1.7 Thesis layout	6
2 Literature survey	8
2.1 Introduction	8
2.2 Fault Detection and Diagnosis in general	8
2.3 Fault Detection and Diagnosis approaches	10
2.3.1 Model-based approaches	10
2.3.1.1 Quantitative	10
2.3.1.2 Qualitative	12
2.3.2 Data-driven approaches	13
2.3.2.1 Quantitative	13
2.3.2.2 Qualitative	13

2.3.3	Hybrid approaches	14
2.3.3.1	General combinations	14
2.3.3.2	Energy-based FDI	14
2.4	Advantages and shortfalls	16
2.5	Performance criteria	17
2.5.1	Patel and Kamrani [55]	17
2.5.2	Venkatasubramanian et al. [2]	18
2.5.3	Reddy [56]	18
2.5.3.1	Reddy's fault detection metrics	20
2.5.3.2	Reddy's fault diagnosis metrics	21
2.5.4	Kurtoglu et al. [57]	21
2.6	Applications	22
2.7	Conclusion	23
3	Gas-to-liquids model	25
3.1	Introduction	25
3.2	Synthetic fuel	25
3.2.1	Historical background	25
3.2.2	General process	28
3.3	Gas-to-liquids	28
3.3.1	Synthesis gas production	29
3.3.1.1	Pre-reforming	29
3.3.1.2	Reforming	29
3.3.1.3	Syngas cleaning and conditioning	29
3.3.2	Fischer-Tropsch synthesis	31
3.3.2.1	The process	31
3.3.2.2	Anderson-Schulz-Flory distribution	31
3.3.3	Product upgrading	33
3.4	Developed GTL model	33
3.4.1	Simulation software	34
3.4.2	Modelling assumptions	34
3.4.2.1	Feedstocks	34
3.4.2.2	Process flow	34
3.4.2.3	Thermodynamic package	35
3.4.2.4	Noise	35
3.4.3	Modelled process at a glance	35
3.4.4	Autothermal reformer	36
3.4.5	Fischer-Tropsch reactor	38
3.4.6	Model validation	40

3.4.7	Recycling	40
3.5	Fault conditions	43
3.5.1	Fault rationale	43
3.5.2	Fault sets	43
3.5.2.1	Fault set F_{1q_r}	45
3.5.2.2	Fault set F_{2q_r}	46
3.5.2.3	Fault set F_{3q_r}	46
3.6	Conclusion	49
4	Energy characterisation	50
4.1	Introduction	50
4.2	Background to exergy	50
4.3	Exergy calculations	51
4.3.1	Reference Environment	51
4.3.2	Total exergy	51
4.3.3	Physical exergy	52
4.3.3.1	Theory	52
4.3.3.2	User variable	52
4.3.3.3	Validation	53
4.3.4	Chemical exergy	53
4.3.4.1	Theory	53
4.3.4.2	User variable	53
4.3.4.3	Validation	54
4.4	Energy characterisation	56
4.5	Conclusion	58
5	Exergy-based fault detection: a threshold approach	59
5.1	Introduction	59
5.2	Methodology	59
5.2.1	Quick overview	59
5.2.2	Threshold approach	60
5.2.3	Assessment metrics	61
5.2.3.1	Detection	62
5.2.3.2	Isolability	62
5.2.3.3	Isolation	63
5.2.3.4	Sensitivity	63
5.2.3.5	Storage and computational requirements	63
5.3	Results	63
5.4	Approach performance evaluation	64
5.4.1	Detection	64

5.4.2	Isolability	67
5.4.3	Isolation	67
5.4.4	Sensitivity	68
5.4.5	Storage and computational requirements	68
5.5	Conclusion	72
6	Energy-based fault detection: a graph matching approach	73
6.1	Introduction	73
6.2	Methodology	74
6.2.1	Quick overview	74
6.2.2	Graph matching approach	74
6.2.3	Assessment metrics	78
6.2.3.1	Detection	78
6.2.3.2	Isolability	79
6.2.3.3	Isolation	79
6.2.3.4	Sensitivity	79
6.2.3.5	Storage and computational requirements	79
6.3	Results	79
6.4	Approach performance evaluation	82
6.4.1	Detection	82
6.4.2	Isolability	82
6.4.3	Isolation	82
6.4.4	Sensitivity	83
6.4.5	Storage and computational requirements	83
6.5	Conclusion	83
7	Energy-based fault detection: eigendecomposition approach	85
7.1	Introduction	85
7.2	Methodology	85
7.2.1	Quick overview	85
7.2.2	Eigendecomposition	86
7.2.2.1	Qualitative	87
7.2.2.2	Quantitative	88
7.2.3	Assessment metrics	90
7.2.3.1	Detection	90
7.2.3.2	Isolability	90
7.2.3.3	Isolation	90
7.2.3.4	Sensitivity	90
7.2.3.5	Storage and computational requirements	90
7.3	Results	90

7.3.1	Approach III.A	91
7.3.2	Approach III.B	93
7.4	Approach performance evaluation	95
7.4.1	Approach III.A	95
7.4.1.1	Detection	95
7.4.1.2	Isolability	95
7.4.1.3	Isolation	96
7.4.1.4	Sensitivity	96
7.4.1.5	Storage and computational requirements	96
7.4.1.6	Summarising remarks	96
7.4.2	Approach III.B	97
7.4.2.1	Detection	97
7.4.2.2	Isolability	98
7.4.2.3	Isolation	98
7.4.2.4	Sensitivity	98
7.4.2.5	Storage and computational requirements	99
7.4.2.6	Summarising remarks	99
7.5	Approaches comparison	99
7.6	Conclusion	103
8	Conclusion	104
8.1	Introduction	104
8.2	Outcome of research objectives	104
8.2.1	GTL model	104
8.2.2	Energy characterisation	104
8.2.3	Fixed-threshold approach - Approach I.A	105
8.2.4	Graph-matching D_C -value - Approach II.B	105
8.2.5	Graph-matching eigenvalues - Approach III.A/B	105
8.2.6	Comparison of approaches	106
8.3	Contribution	106
8.4	Future work and recommendations	107
8.4.1	GTL as benchmark process	107
8.4.2	Approach sensitivity	107
8.4.3	Multiple faults	107
8.4.4	Dynamic system	107
8.4.5	Inclusion of sensor noise	108
8.5	Closure	108
	Bibliography	108

Appendices	124
A Central Composite Rotatable Design	124
A.1 Introduction	124
A.2 Factor identification	125
A.3 Ranges of factors	125
A.4 Design matrix	126
A.5 Mathematical equations	126
A.6 Response evaluation	129
B Standard chemical exergy calculations	131
C HYSYS® user variables	137
C.1 Set-up	137
C.2 VBA code	137
D Normal operating conditions	140
E Calculating the threshold value	144
E.1 Calculation	144
F Eigenvalues' standard deviation	146
F.1 Calculation	146
F.2 Data	147
G IFAC World Congress 2020 article	149

LIST OF FIGURES

Figure 1.1	Visual summary of the thesis' methodology	5
Figure 1.2	Graphical summary of the layout of the thesis document	7
Figure 2.1	A general process monitoring loop [26]	9
Figure 2.2	A generic fault diagnosis framework [2, 16]	10
Figure 2.3	Various FDD techniques categorised [2, 16]	10
Figure 2.4	Graphical representation of a model-based fault diagnosis scheme [30]	11
Figure 2.5	Time line of surveyed energy-based FDI approaches	15
Figure 2.6	Reddy's evaluation procedure for detecting and diagnosing faults [56]	19
Figure 2.7	Confusion matrix of the FDD system outputs versus the true conditions	21
Figure 2.8	Different chemical systems and the prominent FDD techniques found in literature	24
Figure 3.1	A map showing the various global synthesising companies and their affiliations [127]	26
Figure 3.2	Indirect conversion of carbonaceous feedstock to synthetic fuels and chemicals (Adapted [132, 133])	28
Figure 3.3	The three major sections of a GTL process (Adapted [22])	29
Figure 3.4	Probability of chain growth to subsequent hydrocarbons in FT reactions [23, 24]	33
Figure 3.5	An overview of the process flow of the developed GTL process	36
Figure 3.6	HYSYS® process flow diagram of syngas production section with relevant stream numbers	36
Figure 3.7	HYSYS® process flow diagram of Fischer-Tropsch section	39
Figure 3.8	The ASF distribution of the FTR products C ₂ -C ₂₀ (Stream 12)	41
Figure 3.9	The Aspen HYSYS® process flow diagram of the developed GTL process	42
Figure 3.10	The prominent process units seen in a GTL process	43
Figure 3.11	The GTL process showing the considered fault locations	48
Figure 5.1	Graphical representation of the threshold approach (Approach I.A)	60
Figure 5.2	Graphical representation of the applied threshold	61
Figure 5.3	Visual depiction of the cumulative performance metrics of Approach I.A	72
Figure 6.1	A graphical representation of the graph matching approach (Approach II.B)	74
Figure 6.2	The graph of the GTL process showing the nodes, edges and energy attributes	76

Figure 6.3	Visual depiction of the cumulative performance metrics of Approach I.A	84
Figure 7.1	A graphical representation of the eigendecomposition approach (Approach III)	86
Figure 7.2	Visual depiction of the overall performance metrics of Approach III.A	97
Figure 7.3	Visual depiction of the overall performance metrics of Approach III.B	99
Figure 7.4	Graphical representation of the approaches' performance for (a) F_{pq_R} , (b) F_{pq_1} and (c) cumulative	101
Figure A.1	Generalised methodology of CCRD	124
Figure A.2	The response surface plots for the effects on (a) temperature by x_1 and x_2 , (b) composition by x_1 and x_3	130
Figure B.1	Free energies of formation for various hydrocarbons for temperature range 0 - 1500 K [170]	134
Figure B.2	Free energies of formation for various hydrocarbons for temperature range 0 - 1500 K (continued) [170]	135
Figure C.1	The <i>user variable</i> created to calculate the physical exergy	138
Figure C.2	The <i>user variable</i> created to calculate the chemical exergy per phase	139

LIST OF TABLES

Table 2.1	Quantitative model-based methods advantages and shortfalls	16
Table 2.2	Qualitative model-based methods advantages and shortfalls	16
Table 2.3	Data-driven methods advantages and shortfalls	17
Table 2.4	Hybrid methods advantages and shortfalls	17
Table 2.5	Summary of Patel and Kamrani’s assessment criteria [55]	18
Table 2.6	Desirable characteristics of an FDD system as proposed by Venkatasubramanian [2]	19
Table 2.7	A summary of Reddy’s FDD assessment methodology [56]	20
Table 2.8	Assessment metrics as detailed by Kurtoglu et al. [57]	22
Table 3.1	Companies and their plants, years operational, Barrels per day capacity, type of feedstock, and location [24, 127–131]	27
Table 3.2	The various syngas production reforming technologies (Adapted [24, 133, 135])	30
Table 3.3	FT synthesis classification temperatures	31
Table 3.4	Different FT technologies seen commercially over the years [24, 127–130]	32
Table 3.5	Syncrude-to-product conversions and product details	34
Table 3.6	Feed ratios of the ATR components	37
Table 3.7	Syngas production section stream information as simulated in HYSYS®	37
Table 3.8	Simulated molar fractions of the main components of the syngas (Stream 5)	37
Table 3.9	The stoichiometric coefficients of the CO consumption (Equation (3.7))	39
Table 3.10	The kinetic values and units used for the reactions in HYSYS®	39
Table 3.11	Fischer-Tropsch synthesis stream information as simulated in HYSYS®	40
Table 3.12	Summary of the weight fraction per carbon number (C_{1-14})	40
Table 3.13	Summary of the weight fraction per carbon number (C_{15-20})	41
Table 3.14	Stream information of the simulated GTL process	41
Table 3.15	Common recurring causes and effects of unit failures [1, 148]	44
Table 3.16	Percentage magnitudes of signifier r	45
Table 3.17	The location and details of simulated faults F_{1q_r}	45
Table 3.18	The location and details of simulated faults F_{2q_r}	46
Table 3.19	The location and details of simulated faults F_{3q_r}	47

Table 4.1	Hand calculated and user variable exergy values for the methane stream (Stream 1) compared	53
Table 4.2	Hand calculated chemical exergy for components in syngas stream (Stream 5)	55
Table 4.3	Hand calculated and user variable exergy values for the syngas stream (Stream 5) compared	55
Table 4.4	Physical exergy, chemical exergy, and energy flow of each stream of the NOC	57
Table 4.5	Important fault datasets used in this study	57
Table 5.1	Confusion matrix and relevant detection rates calculations	62
Table 5.2	Threshold approach methodology outputs per stream for normal condition Normal ₁	64
Table 5.3	The qualitative matrices of dataset F_{pq_R} after applying the threshold function	65
Table 5.4	The qualitative matrices of dataset F_{pq_1} after applying the threshold function	66
Table 5.5	Confusion matrix when applying Approach I.A on dataset (a) F_{pq_R} and (b) F_{pq_1}	67
Table 5.6	The isolability performance of Approach I.A	67
Table 5.7	The isolation performance of Approach I.A	68
Table 5.8	A summary of the performance metrics for Approach I.A	68
Table 5.9	Detection, isolability, and isolation metrics of dataset F_{pq_R} compared to itself	69
Table 5.10	Detection, isolability, and isolation metrics of dataset F_{pq_1} compared to itself	70
Table 5.11	Detection, isolability, and isolation metrics of dataset F_{pq_R} compared to F_{pq_1}	71
Table 6.1	The corresponding process units and streams used to construct the GTL graph	76
Table 6.2	Detectability, isolability, and isolation of fault dataset F_{pq_R}	80
Table 6.3	Detectability, isolability, and isolation of fault dataset F_{pq_1}	81
Table 6.4	Confusion matrix when applying Approach II.B on dataset (a) F_{pq_R} and (b) F_{pq_1}	82
Table 6.5	The isolability performance of Approach II.B	82
Table 6.6	The isolation performance of Approach II.B	83
Table 6.7	A summary of the performance metrics for Approach II.B	83
Table 7.1	Qualitative assignments using fault condition F_{137} as operational example	88
Table 7.2	Example of quantitative differences of fault condition F_{213}	89
Table 7.3	Detectability, isolability and isolation of fault dataset F_{pq_R} when applying Approach III.A	91
Table 7.4	Detectability, isolability and isolation of fault dataset F_{pq_1} when applying Approach III.A	92
Table 7.5	Detectability, isolability and isolation of fault dataset F_{pq_R} when applying Approach III.B	93
Table 7.6	Detectability, isolability and isolation of fault dataset F_{pq_1} when applying Approach III.B	94
Table 7.7	Confusion matrix when applying Approach III.A to dataset (a) F_{pq_R} and (b) F_{pq_1}	95
Table 7.8	The isolability performance of Approach III.A	95

Table 7.9	The isolation performance of Approach III.A	96
Table 7.10	Comparison of the performance metrics of dataset F_{pq_1} for Approach II.B and Approach III.A	96
Table 7.11	A summary of the performance metrics for Approach III.A	97
Table 7.12	Confusion matrix when applying Approach III.B to dataset (a) F_{pq_R} and (b) F_{pq_1} . . .	97
Table 7.13	Comparison of the overall detection metrics for Approach III.A and Approach III.B . .	98
Table 7.14	Comparison of isolability performance of Approach III.A and Approach III.B	98
Table 7.15	Comparison of isolation performance of Approach III.A and Approach III.B	98
Table 7.16	Comparison of the performance metrics of dataset F_{pq_1} for Approach III.A and Approach III.B	99
Table 7.17	A summary of the performance metrics for Approach III.B	99
Table 7.18	A summary of the performance metrics of the various approaches investigated	100
Table 7.19	Visual summary of the detection, isolability, and isolation of F_{pq_R} of all approaches .	102
Table 7.20	Visual summary of the detection, isolability, and isolation of F_{pq_1} of all approaches .	102
Table 8.1	Summary of investigated energy-based FDI approaches and their details	104
Table A.1	Summary of the factors and response variables used for the CCRD	125
Table A.2	The calculated g_{x_i} and t_{x_i} values for the three factors	126
Table A.3	The actual and coded values of the three factors	126
Table A.4	The design matrix along with the two response variables' simulation values	128
Table A.5	Assumed and eventual values for the chosen feed ratios	130
Table B.1	(a) Hypothetical chamber showing fuel conversion and (b) the corresponding combustion equation	131
Table B.2	Gibbs of formation and standard chemical exergy values used to calculate the hydrocarbons	132
Table B.3	Comparison between known and calculated standard chemical exergy	132
Table B.4	Gibbs function of formation and calculated standard chemical exergy $\text{CH}_4 - \text{C}_{11}\text{H}_{24}$.	132
Table B.5	Gibbs function of formation and calculated standard chemical exergy of $\text{C}_{12}\text{H}_{26} - \text{C}_{30}\text{H}_{62}$	133
Table B.6	Standard chemical exergy of the other substances included in simulation [25]	133
Table B.7	Gibbs function of formation ($\bar{g}^0$) for some common substances	136
Table C.1	User variable option setup	137
Table D.1	Average physical exergy, chemical exergy, and energy flow making up NOC	140
Table D.2	The recorded physical exergy data for the 10 runs and resultant average (NOC)	141
Table D.3	The recorded chemical exergy data for the 10 runs and resultant average (NOC)	142
Table D.4	The recorded energy flow data for the 10 runs and resultant average (NOC)	143

Table E.1	Parameters and the formulae used to quantify the error percentage of the simulation variations	144
Table E.2	Calculating the threshold value κ by using the simulation variations	145
Table F.1	Eigenvalues of cost matrices used to determine the standard deviation	147
Table F.2	Calculated standard deviation for each eigenvalue	148

ABBREVIATIONS

ANN	Artificial Neural Network
ASF	Anderson-Schulz-Flory
ATR	Autothermal Reformer
CCRD	Centrally Composite Rotatable Design
CD	Correct diagnosis
CPO	Catalytic partial oxidation
CSTR	Continuous stirred-tank reactor
DCL	Direct coal liquefaction
DEs	Differential equations
ED	Experimental Design
EKF	Extended Kalman Filter
EOS	Equation of State
FCCU	Fluid catalytic cracking unit
FDD	Fault Detection and Diagnosis
FDI	Fault Detection and Isolation
FN	False negative
FP	False positive
FT	Fischer-Tropsch
FTR	Fischer-Tropsch reactor
GLRT	Generalised likelihood ratio test
GTL	Gas-to-liquids
HEOM	Heterogeneous Euclidean Overlap Metric
HER	Heat exchange reforming
HTFT	High-temperature Fischer-Tropsch
HVAC&R	Heating, ventilating, air-conditioning, and refrigerating
KF	Kalman Filter

LTFT	Low-temperature Fischer-Tropsch
MD	Misdiagnosis
MTFB	Multi-tubular fixed bed
MTFT	Medium-temperature Fischer-Tropsch
ND	No detection
ND_G	No diagnosis
NOC	normal operating condition
NUIOs	Non-linear Unknown Input Observers
ODEs	Ordinary differential equations
OLS	Ordinary least square
PCA	Principal Component Analysis
PCEG	Possible cause and effect graph
PCI	Petrochemical industry
PFR	Plug flow reactor
PLS	Partial Least Square
POX	Partial Oxidation
PR	Peng-Robinson
QRR	Qualitative Redundant Relation
QTA	Qualitative Trend Analysis
RE	Reference environment
RS	Response surface
Sasol	South African Coal and Oil Company
SDG	Signed digraph
SMR	Steam methane reforming
SPC	Statistical Process Control
SPM	Statistical process monitoring
TEP	Tennessee Eastman Process
TN	True negative
TP	True positive
UIOs	Unknown Input Observers
VBA	Visual Basic for Applications
XTL	Feed-to-liquid

NOMENCLATURE

b_{ch}	Intrinsic chemical exergy
B_{ch}	Total chemical exergy
b_{ch}^0	Standard chemical exergy
b_{kin}	Intrinsic kinetic exergy
b_{ph}	Intrinsic physical exergy
B_{ph}	Total physical exergy
b_{pot}	Intrinsic potential exergy
b_{tot}	Intrinsic total exergy
C	Cost matrix
D_c	Distance value of cost matrix
e	Effort
$\dot{E}$	Energy flow
E	Edges (or links or arcs)
$e\%$	Percentage error
f	Flow
F_f^0	Free energies of formation
$\bar{g}^0$	Gibbs function of formation
G	Graph
h	Enthalpy
h_0	Reference environment enthalpy
m	Number of samples
μ	Average
n	Total molar flow rate
p	Statistical significance
P	Pressure
P_0	Reference environment pressure

r_{FN}	False negative rate
r_{FP}	False positive rate
r_{TP}	True positive rate
s	Entropy
σ	Standard deviation
s_0	Reference environment entropy
t_{m-1}	t-value
T	Temperature
T_0	Reference environment temperature
V	Vertices (or nodes)
y	Threshold fault element
$\Psi_{(\text{H}_2/\text{CO})}$	Syngas composition response variable
$\Psi_{(T)}$	Syngas temperature response variable
z	Normalised exergy value

CHAPTER 1

INTRODUCTION

1.1 MOTIVATION

Process monitoring is deemed a vital aspect in ensuring the consistent and efficient operation of industrial process plants. Traditionally, operators are tasked with supervision of a plant's operations and health. When anomalies occur within the system, the operators are expected to detect, diagnose and rectify the phenomenon promptly. With the advances seen within various technological avenues, plants are occasionally upgraded, resulting in even more complex processes having to be monitored. This is especially true for chemical process plants that include recycling streams and control systems that could potentially conceal the effects of faults. Consequently, the responsibilities and responses required of the operators could easily escalate beyond their capabilities. Mishandling of such events could have costly repercussions, not only risking human life and the environment but also having considerable financial implications. A well-known incident that illustrates this is the methyl isocyanate (MIC) leak in Bhopal, India which claimed thousands of lives, as stated by Kletz [1]. Another costly incident, according to [2], was the Kuwait Petrochemical Mina al Ahmadi oil refinery explosion which resulted in \$100 million in damages. This is where Fault Detection and Diagnosis (FDD) is of interest. FDD, which makes up the largest part of process monitoring procedures, endeavours to minimise the impact of faults by substituting the efforts required of operators with an automatic computerised system. An FDD scheme uses variables observed within the system to determine whether a fault is present and to - ideally - diagnose the exact details of the said fault (location, magnitude, root cause, etc.).

Usually, FDD approaches are classified as being either model-based or data-driven, with the main distinction being the availability of an analytical model. Data-driven methods require large amounts of data and processing, without relying on a priori knowledge or analytical models. Excluding qualitative knowledge-based Fault Detection and Isolation (FDI) systems (expert systems) [3], the remaining data-driven approaches can be categorised as being either statistical in nature or based on machine learning approaches. Data-driven methods are especially dominant within the petrochemical industry (PCI). Qin et al. [4] applied several statistical process monitoring (SPM) schemes to an industrial polymer film process. Other studies applied principal component analysis (PCA) techniques to continuous stirred tank

reactors (CSTRs) [3, 5] and plug-flow reactors [6]. The use of mathematical models to generate data is also often seen [3, 7]. Though it would not be impossible to develop, analytical models for PCI plants are regarded as quite challenging. Tidriri et al. [8] clearly illustrate the associated complexity of developing an analytical model of the popular Tennessee Eastman Process (TEP). In a review series done by Gao et al. [3, 9] signal-based diagnosis is introduced, which makes use of measurement signals rather than analytical models. Signal-based diagnosis is still considered to be model-based since the normal signal is known a priori. A key advantage of model-based FDI is that the model can be used to mathematically prove the extent of the technique. As such, model-based FDI would be ideal for volatile applications such as petrochemical processes. Given that data-driven methods are dominant within the PCI and the advantages afforded by model-based techniques, the development of a model-based FDI approaches for the PCIs is warranted. Multiple authors have suggested signal-based FDI that measures the energy flows within a process plant. Du Rand et al. [10] applied an entropy-enthalpy approach to the Brayton cycle of a nuclear power plant, and Marais et al. [11] showed that condition monitoring of a CSTR from an energy perspective could be achieved. Marais [12] later shows that exergy can successfully be utilised to perform FDI for an autothermal reformer (ATR).

Since it is evident that no single FDI scheme will perform flawlessly in all circumstances [13], the development of hybrid approaches is required. Hybrid FDI schemes usually combine two or more existing techniques to exploit the various advantages offered by the constituent techniques. The work done in [14] depicts a typical example of a hybrid FDI approach which is a combination of statistical and machine learning techniques. Combinations of signal- and observer-based techniques have also proven to be successful when applied to a classical two-tank system [15]. According to [13], one won't necessarily obtain better results by simply combining random techniques. Based on conjecture, a combination of model-based and data-driven FDI schemes should provide significant improvements in terms of both FDI performance and reliability. In [16] the application of the exergy-based FDI scheme (classified as being a hybrid scheme) to a process containing recycle streams was identified as an area of future work. Seeing that the ATR plant used by [16] did not include any recycle streams, the applicability of the exergy-based approach to petrochemical plants that include recycle streams remains unknown. Continuing with this motif, further work done by [17] made use of graph matching and eigenvalues as a means to energy-visualisation to achieve FDI for a counter-flow heat exchanger. Energy-visualisation refers to the manner in which energy properties are packaged into an attributed graph to retain physical, structural information. More recently, Naser [18] and Uren et al. [19] expanded upon the work done by [17]. Naser [18] shows how various energy-visualisation approaches could be used within a Brayton cycle power conversion unit which is seen as a complete thermodynamic system. In the article by Uren et al. [19], the technique was applied to a heated two-tank system. Based on these researchers' promising findings and identified open questions, and given the limited examples of model-based FDI in the PCIs, the limitations, suitability, and performance of existing implementations should be explored. The subsequent sections introduce the reader to the research methodology and contributions. Some of the research outputs are also outlined. Lastly, the document's layout is outline per chapter and related content.

1.2 PROBLEM STATEMENT

With the continuous expansion of the FDD field, hybrid approaches seem to be at the forefront. Keeping with this theme, the study will evaluate the feasibility and performance of energy-based approaches as a means to Fault Detection and Isolation (FDI) of a gas-to-liquids (GTL) process. The main focus of the investigated approaches will be the hybridisation of energy properties and structural information.

1.3 RESEARCH AIMS AND OBJECTIVES

The following lists the main research objectives for this study:

- Develop a steady-state simulation model of a gas-to-liquids (GTL) process.
- Characterise the normal and faulty GTL process behaviours in terms of energy.
- Evaluate the FDI capabilities of an exergy-based fixed-threshold approach when applied to a petrochemical process (the GTL process) that includes a recycle stream.
- Determine the feasibility of applying energy-based, graph-based approaches as a means to FDI of a GTL process.
- Compare the FDI capabilities of the various approaches.

1.4 RESEARCH METHODOLOGY

After surveying and presenting the relevant literature concerning Fault Detection and Diagnosis, the following efforts and tasks were of cardinal importance. A graphical representation of the endeavour is shown in Figure 1.1, where each chronological facet corresponds to a numbered circle within the diagram.

① The first requirement of this project is to have a working and representative simulation model of a GTL process. In order to achieve such a model, the literature and knowledge pertaining to modelling a GTL plant are deemed crucial. As with many other industrial plants, some information is obscured because of its proprietary nature. The five most descriptive resources that are of great value are the works of [16, 20–23]. Some additional research revealed a dissertation with a detailed model of a GTL process [24]. Utilising the information given in these six resources, the GTL operating conditions, elements and components as well as the expected product distribution can be established. The GTL process will be simulated as a steady-state model making use of a commercial process simulator called, Aspen HYSYS®. ② To validate the simulated products that are produced, the product distribution will be compared to the theoretical distribution found in literature.

③ The third phase of the project entails characterising the modelled GTL in terms of energy (exergy and energy flow). To get a handle on the terminology and concepts of exergy, some chapters in [25] are

studied. To automatically calculate the exergy within the HYSYS[®] model, *user variables* will be created. *User variables* are snippets of Visual Basic for Applications (VBA) code that is used to manipulate various components within HYSYS[®] [16]. Specific attention will be given to calculating the chemical exergy of streams that consist of multiple phases (gas, light liquid and heavy liquid). (4) It is imperative to validate the exergy quantities obtained from the *user variables*. To do so, the calculated exergy values will be compared to hand calculations under the same operating conditions.

(5) Naturally, with the field of study being FDI, fault conditions need to be defined. The book by Kletz [1] is studied to determine typical component/part failures and the effect such faults would have on a system. After relevant faults are identified, the (6a) normal and (6b) faulty operating conditions will be simulated individually in order to obtain and (7) record the desired energy data.

The first energy-based FDI approach to assess is the technique proposed in the work of Marais [16]. To apply the approach, (8) a threshold value will need to be established first. The threshold value is based on the simulation variation seen within HYSYS[®]. After employing the approach, the (9) detection, isolability, and isolation performance will be calculated and evaluated accordingly. The next two approaches examined in this study are based on graph matching theoretical aspects. For both of the graph-based approaches being investigated, the (10) development of a database is required. The premise of the approaches is that a database is loaded with energy graphs; containing a normal operating condition (NOC) and faulty conditions. The approaches will then compare an unknown operational condition's energy graph to those stored within the database. For the first graph-based approach, the dissimilarities, (11) quantified by calculating a distance parameter D_C , will indicate the likeliest match within the database. (12) Subsequently, signifying a fault-free or faulty operational condition. (13) For the second graph-based approach, the groundwork remains the same, but the matching mechanisms are based on changes seen in the eigenvalues. (14) Once again, the fault detection, isolability, and isolation performance will be assessed. Finally, an overall comparison and conclusion will be drawn of the various approaches covered in this study.

1.5 CONTRIBUTION OF RESEARCH

The contribution of the research lies in the modification, application, and evaluation of existing energy-based FDI approaches to a petrochemical process not yet considered within the FDD field. The work done by Marais [16], van Graan [17], Neser [18], and Uren et al. [19] show the suitability of various energy-based FDI approaches as applied to their respective systems. These studies qualitatively discuss some FDI performance properties, but do not specifically compare their performance definitively. The main focus is, therefore, to determine - after some alterations - the applicability and performance of some of the previously developed energy-based approaches, especially compared to one another, when applied to a single, larger-scale system (GTL process).

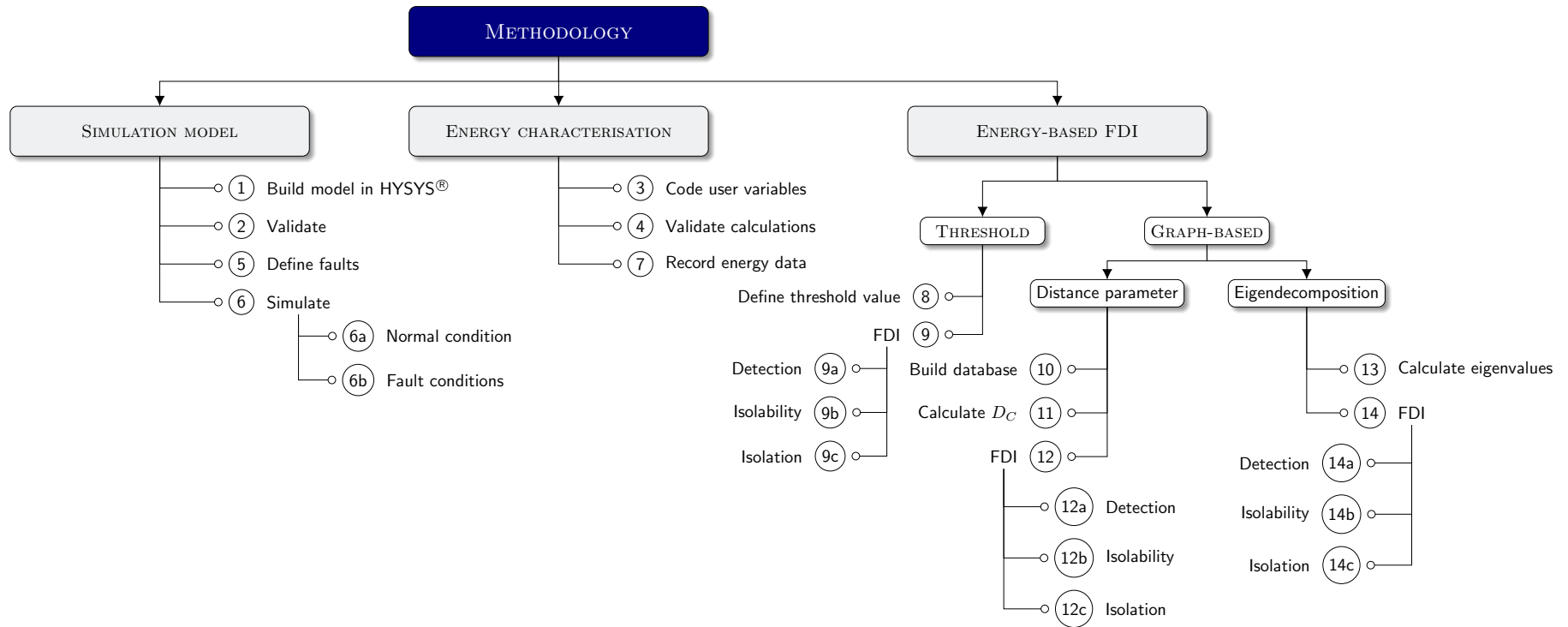


Figure 1.1: Visual summary of the thesis' methodology

1.6 PUBLICATIONS FROM THE RESEARCH

The following publications originated from this study. The first publication listed is available on-line (open access). The second article was presented virtually at the IFAC World Congress. The conference proceedings are not available as of yet; therefore, the article is given in Appendix G. The last journal article is in progress and will be submitted in January 2021.

- S. Greyling, H. Marais, G. van Schoor, and K.R. Uren, “Application of exergy-based fault detection in a gas-to-liquids process plant,” *Entropy*, vol. 21, no. 6, p. 565, 2019, (available at <https://doi.org/10.3390/e21060565>).
- S. Greyling, G. van Schoor, K.R. Uren, and H. Marais, “Exergy graph-based fault detection and isolation of a gas-to-liquids process,” *IFAC-PapersOnLine: Proceedings of the 21st IFAC World Congress*, 2020, (included in Appendix G).
- S. Greyling, G. van Schoor, K.R. Uren, and H. Marais “Comparative study on energy-graph based fault detection and isolation techniques applied to a gas-to-liquids process,” *Computers and Chemical Engineering*, (To submit in January 2021).

1.7 THESIS LAYOUT

The complete thesis document consists of 8 chapters. A graphical summary of the chapters and their content is given in Figure 1.2. **Chapter 2** is used to survey relevant literature found in the Fault Detection and Diagnosis (FDD) field, with special attention being given to the existing energy-based FDI. As this study utilises a gas-to-liquids (GTL) process, **Chapter 3** is used to first provide some historical and operational information of the process, with the proceeding sections giving specific details surrounding the modelling and validation of the simulation model. The FDI techniques being investigated in this study make use of exergy and energy flow data. Thus, the exergy calculations and energy characterisation of the GTL process are discussed comprehensively in **Chapter 4**. The first FDI approach applied to the GTL process is the exergy-based threshold approach proposed by Marais [16]. The methodology, results, and assessment thereof is documented in **Chapter 5**. **Chapter 6** looks at the first energy-based, graph-based FDI approach which utilises a distance parameter (D_C). The methodology, results, and assessment are discussed in a similar manner to **Chapter 5** to ensure that findings can easily be compared. The third and final approach is explored in **Chapter 7**, which makes use of energy-based eigendecomposition as a means to FDI. The approaches and their performance metrics - one of the important outcomes of the study - are compared in the last section of the chapter. **Chapter 8** concludes the thesis by recapitulating some findings and discussing the contributions made.

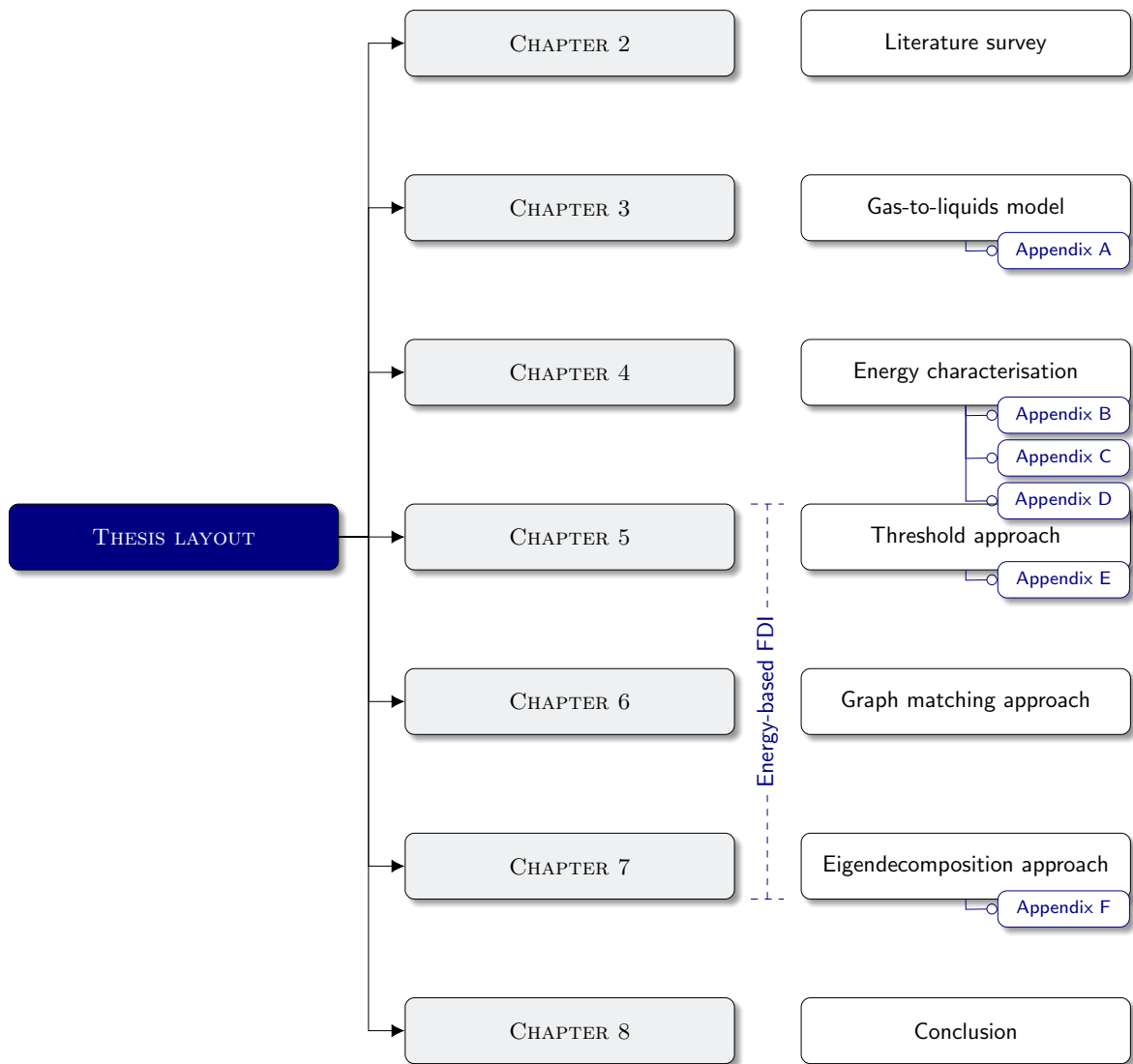


Figure 1.2: Graphical summary of the layout of the thesis document

CHAPTER 2

LITERATURE SURVEY

2.1 INTRODUCTION

As this study investigates energy-based FDI approaches, this chapter is used to survey the classic process monitoring literature. The survey starts by looking at the broader Fault Detection and Diagnosis (FDD) field, relevant terminology, and approach-categories seen. Each category and constituent techniques are defined and briefly discussed. The existing energy-based FDI approaches are deemed of great importance and are thus reviewed in more detail. Additionally, some advantages and shortfalls of the different FDI approaches are summarised. To eventually assess an FDI technique's performance, the criteria generally seen within literature are scrutinised. Lastly, a survey of popular applications (systems) is conducted.

2.2 FAULT DETECTION AND DIAGNOSIS IN GENERAL

Process monitoring is an essential aspect in ensuring the consistent and efficient operation of plants. Should a prohibited anomaly occur within the plant, operators are traditionally expected to detect, diagnose and rectify the situation in the shortest time possible. When events (or faults) are not adequately dealt with, the effects thereof usually result in costly incidents [2, 13, 26]. Such events not only impact the plant's productivity but risk the safety of human life, the environment and the company's finances. This is where Fault Detection and Diagnosis (FDD) can be advantageous. An FDD scheme is an automatic computerised system which aids in the detection and diagnosis of a fault within a system, effectively minimising the reliance upon operators and improving the efficacy in with which the event is handled. The formal definitions of the following terms should be noted [2, 26, 27]:

- A *disturbance* is an unknown, uncontrollable input which acts in on the system.
- A *fault* is defined as any deviation of an observed or calculated process parameter from its acceptable operating range.
- A *failure* is a permanent interruption of a system's ability to perform a required function and can be seen as a *root cause* or *basic event*.

A fault generally falls into one of the subsequent classifications [2, 16, 28]:

- *Additive process faults* are defined as unknown inputs acting on the plant which cause changes in the plant outputs independent of the known inputs. This usually signifies a structural change such as leaking pipes or stuck valves.
- *Multiplicative process faults* are sudden or graduate changes in plant parameters. The plant outputs are changed and are dependent on the magnitude of the known inputs. Faults like these include the deterioration of plant equipment or variations in reactants.
- *Sensor faults* are the differences between measured and actual values of the plant variables.
- *Actuator faults* are the variations found between the input command of the actuator and its true output.

A general process monitoring loop is depicted in Figure 2.1. The *detection* facet determines whether a fault is present within the system. The *diagnosis* aspect is usually a grouping of *isolation* and *identification*; where *isolation* is finding the exact location of the fault and *identification* describing the magnitude of the fault. In most cases, fault *detection* is crucial, with the fault *isolation* being just as important. It is seldom that the fault *identification* outcomes warrant the extra effort required. For this reason, most systems only include the first two tasks and are referred to as Fault Detection and Isolation (FDI). The fourth element, the *recovery*, is concerned with the rectification of the fault but is not included in FDD schemes [26, 28–30].

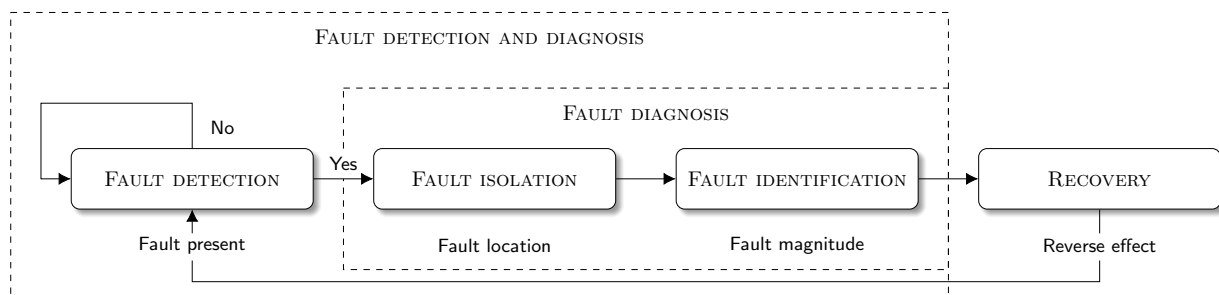


Figure 2.1: A general process monitoring loop [26]

As with most research fields, FDD has certain conventions which are customarily adopted: (1) According to [28], the system is assumed to have no faults present initially, with the fault appearing only some unknown time later. (2) There is a subjective distinction between faults and disturbances. Although both are deterministic and unknown inputs to the system; faults are seen as events that need to be detected and isolated, whereas disturbances are occurrences that should ideally be negligible. (3) Unstructured uncertainties, process noise, and measurement noise are not included in the scope of the FDD system. The unstructured uncertainties emanate from faults that were not modelled a priori. Here, process noise refers to the discrepancies between the model parameters and the actual system. The measurement noise is ascribed to high-frequency additive elements in sensor measurements [2]. Figure 2.2 shows the various components of a fault diagnostic framework of a controlled process system. It also shows the different origins of the mentioned failures.

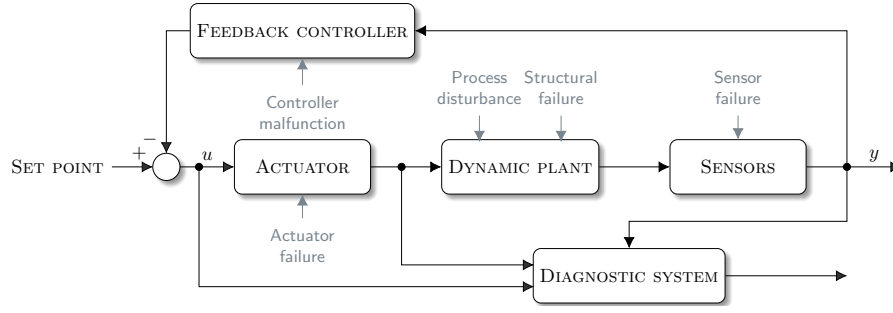


Figure 2.2: A generic fault diagnosis framework [2, 16]

2.3 FAULT DETECTION AND DIAGNOSIS APPROACHES

Since FDD research started in the 1970s, the field of process monitoring has grown extensively to include many different techniques and applications. Based on the nature of the knowledge utilised, the techniques found in literature are categorised as being model-based or data-driven. With each approach and technique having its particular strengths and drawbacks, recent works sought to combine the advantages of more than one technique to develop a hybrid approach [2, 26]. Figure 2.3 shows a diagram that summarises the prominent techniques found in literature under the appropriate category. Under each of the categories, the techniques can further be broken down into qualitative and quantitative approaches. The subsequent sections will briefly detail some of the depicted techniques. For a more in-depth discussion, the interested reader is referred to [2, 31–33].

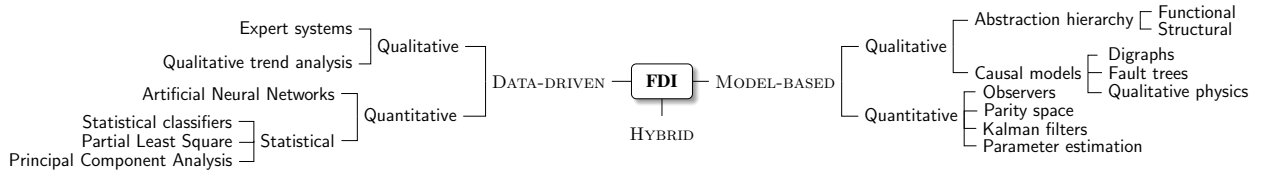


Figure 2.3: Various FDD techniques categorised [2, 16]

2.3.1 MODEL-BASED APPROACHES

Model-based approaches make use of explicit first-principle and fundamental knowledge of a system to develop a process model. Usually, these models encapsulate the normal as well as the faulty behaviours of the systems [30, 33]. To detect a fault generated residuals are evaluated. These residuals are descriptions of the differences between measured process variables and their estimates. A graphical representation of this is shown in Figure 2.4. The development of such residuals can be achieved by a diverse range of methods and are categorised as being either quantitative or qualitative [2, 13].

2.3.1.1 QUANTITATIVE

OBSERVERS

Observers (also referred to as diagnostic observers) are models of systems which are run in parallel to the considered processes. The model calculates the estimates of the process and generates residuals based on

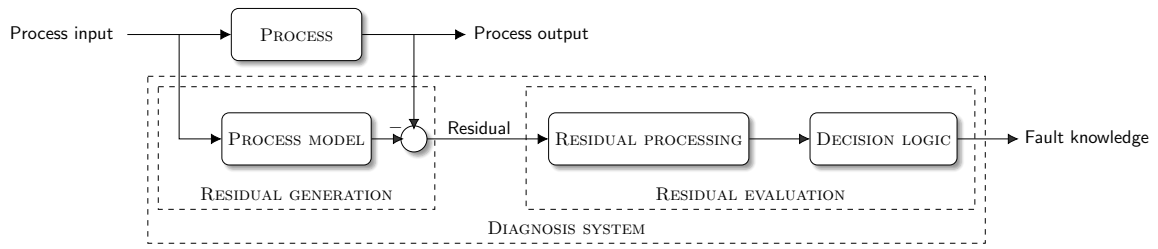


Figure 2.4: Graphical representation of a model-based fault diagnosis scheme [30]

the differences between these estimates and actual measured values [34]. If the generated residuals are close to zero, the system is operating normally. If a significant residual is generated, a faulty condition is indicated. One particular concern with observers is the dependency of the model on the process parameters. In some systems not all process parameters are always known. This directly influences the robustness of the technique. There have been some compelling results of improving the robustness of observers, mostly where disturbances and uncertainties were an issue, by the development of Unknown Input Observers (UIOs) and Non-linear Unknown Input Observers (NUIOs) [35].

PARITY SPACE

Parity space is based on the transformation of the state-space model of the system in order to obtain the parity relations. The benefit of developing parity relations is to attain equations which only depend on measured or known variables; namely the inputs and the outputs. To acquire these parity relations, redundancies between the various variables of the system are used. The shortcomings of the approach are that it does not typically account for significant model uncertainties, unmodelled disturbances or multiplicative faults. Considering that a comprehensive process model is required the application of the technique is essentially limited to linear time-invariant systems and not a plausible task for complex and non-linear systems [13].

KALMAN FILTERS (KF)

If a system's diagnostic model is developed in the space of object states, with the connection between input signal, noise and output signal given by differential/difference equations; a Kalman Filter (KF) is a technique that searches for an optimum estimate using the least-square method [36]. A weakness of the Kalman Filter is its linear characteristics [16]. There have been studies that used Extended Kalman Filters (EKF) to assist with non-linear systems. Nonetheless, model uncertainties and disturbances still adversely influenced the effectiveness of the technique [35].

PARAMETER ESTIMATION

Parameter estimation methods were some of the first studied for performing FDD within real-time industrial systems [37]. These methods rely on the principle that possible faults in an observed system can be linked with specific parameters and states of the mathematical process model of the system. This is then described as an input-output relation [38]. The computational simplicity of these methods is seen as the primary advantage, with it being well-suited for detecting multiplicative faults [37,39]. The major drawback, however, is their poor robustness against external disturbances on the system behaviour; highly accurate estimations become excessively time-consuming [37].

BOND GRAPHS

Bond graphs are graphical representations of physical dynamical systems and the energy transfer between the constituents. The nodes of a bond represent subsystems, junctions or components. The edges are called bonds, and each bond denotes the instantaneous power transfer between nodes. The bonds have two power variables affiliated to them; the effort (e) and flow (f) with $e \times f = \text{power}$. Depending on the considered system, the effort would be the intensive variables such as chemical potential, electrical potential, force, pressure, temperature, and torque. Flow would be the derivatives of extensive variables like current, entropy flow, molar flow, velocity and, volume flow. The causality information encapsulated by bond graphs is said to be one of its most valued properties [33,40].

2.3.1.2 QUALITATIVE

ABSTRACTION HIERARCHY

Abstraction hierarchy is based on the decomposition of a system. The decomposition can be either functional or structural. The overall system's behaviour is inferred by looking at the laws governing its comprising subsystems. Functional hierarchy characterises the means-end relationships between the system and its subsystems, whereas the structural hierarchy depicts the connectivity of the system and its subsystem [31,41].

CAUSAL MODELS

Causal models are depictions of a system's cause and effect behaviours, specifically evaluating these under normal and abnormal conditions [42]. According to [2], three techniques fall under causal models; these being digraphs, fault trees and qualitative physics

DIGRAPHS

Digraphs (also called directed graphs) are graphs of which the nodes represent events or variables of a system, and the directed edges show the relationship between the nodes. Signed digraphs (SDG) are extensions of digraphs with their structure being similar to digraphs. The main differences being (1) that their directed edges are ascribed a positive or negative sign, indicating the proportionality of the effect. (2) The nodes acquire qualitative values (0, + or -) in regards to the node's reference value [31,33].

FAULT TREES

In some processes, the relationship between faults and symptoms are known to a degree. This a priori knowledge can be described in causal relationships: $\text{fault} \rightarrow \text{events} \rightarrow \text{symptoms}$. These qualitative observations are then conveyed in terms of rules which eventually leads to a condition-classification. Traditionally, fault trees are evaluated as binary variables and make use of Boolean equations [39,43].

QUALITATIVE PHYSICS

Qualitative physics or common-sense models use qualitative terms to estimate and describe the behaviour of a system. It endeavours to preserve most of the crucial characteristics and causality without having to implement the relevant mathematics [44]. There are two known approaches within literature; the first is the formulation of qualitative confluence equations based on the ordinary differential equations (ODEs) which dictates the behaviour of the process. These equations can then be solved using qualitative algebra to describe the qualitative behaviour of the system. The second approach involves deriving qualitative behaviours from

the ODEs. These qualitative descriptions of various failures can then be used as a source of knowledge [31,41].

2.3.2 DATA-DRIVEN APPROACHES

Data-driven or process history approaches are the opposite of model-based approaches. As previously stated, model-based approaches make use of a priori knowledge, whereas data-driven approaches need large sets of historical data of the considered system. There exist techniques that transform the data into a priori knowledge; this conversion is known as feature extraction. As with model-based, the transformation can be qualitative or quantitative [32].

2.3.2.1 QUANTITATIVE

ARTIFICIAL NEURAL NETWORKS (ANN)

ANNs are implemented in many different fields, ranging from predictions and classification problems even to pattern recognition. The classification problem is condensed to the determination of the connection weights. These weights being learned and modified by using the discrepancies between the actual and estimated outputs. An ANN's performance largely depends on the chosen structure, learning algorithm and transfer function [13,32]. Literature shows that ANNs deliver excellent detection performance; its biggest drawback being the lack of explanation facilities [16].

PRINCIPAL COMPONENT ANALYSIS (PCA)

Principal Component Analysis is one of the most popular approaches employed within FDD. PCA is a reduction technique (linear dimensionality) that produces lower dimensional representations of the original data whilst still retaining maximum variance. It also describes the correlation between process parameters. The reduction is accomplished by keeping systematic variations and removing any random variations. The PCA approach is, however, linear and needs to be extended accordingly to address non-linear systems [13,26].

PARTIAL LEAST SQUARE (PLS)

Partial Least Square (or latent structure projection) is another linear dimensionality reduction technique. It maximises the covariance between input and output data in the reduced space [13]. According to [26] the nature of plants, the lack of dynamic models of them and the data-availability are reasons why PCA and PLS will continue to dominate.

2.3.2.2 QUALITATIVE

EXPERT SYSTEMS

Expert systems are computer-based applications that are developed by making use of collected knowledge from experts in certain domains. The information is then loaded into a database in the form of if-then-else statements. An inference engine, a software component of expert systems, is responsible for inferring a diagnosis and any additional information needed to reach a conclusion. Some of the main strengths of expert systems are the ease of development and transparent explanation facilities. Nevertheless, the application of the approach is limited to the specific system and knowledge that was initially assimilated [41].

QUALITATIVE TREND ANALYSIS (QTA) Qualitative trend analysis is an approach which abstracts and interprets characteristics and trends from process data. When employing QTA, filtering is seen as an important aspect, as noise could have a significant effect on the results [32,45].

2.3.3 HYBRID APPROACHES

2.3.3.1 GENERAL COMBINATIONS

Researchers have started addressing the shortfalls that single FDD techniques exhibit by combining various FDD techniques. When a combination of techniques is considered, a hybrid approach is obtained. The hybridisation attempts to utilise different types of knowledge, data, and structures to improve the performance of existing schemes [13,26]. These improvements ensure more reliable FDD schemes which are equipped to handle copious data and uncertainties better. Many different combinations are seen in literature, although too many to discuss individually; three main aspects are of importance. Firstly, hybridisation usually focusses on capturing cause and effect information whilst utilising available historical data and complementary data-driven techniques [46–48]. Secondly, some studies have demonstrated that a well-designed hybrid approach holds more advantages than individual techniques [49]. Lastly, one cannot merely combine arbitrary techniques. The resultant combinational FDD scheme should illustrate superior performance and reliability [13].

2.3.3.2 ENERGY-BASED FDI

In recent years, a few different energy-based FDI methods were proposed and developed. Figure 2.5 depicts a time line summarising the energy-based approaches, the corresponding authors, and the considered systems (applications). When surveying these works, two distinct methodologies are seen. On the one side, the techniques proposed by Berton et al. [50], Theilliol et al. [51], and Chen [52] are based on energy conservation fundamentals in order to acquire an energy-balance of the system. Residuals were then obtained by employing parameter estimations and parity space approaches. Chen’s [52] technique looks at utilising an energy-balance that contains stored, dissipated and supplied energies. Fault detection was accomplished by checking the validity of the energy-balance, whereas isolation could be determined from the energy-balances of the subsystems. An advantage highlighted in this study was that one could distinguish fault locations as being either in energy-dissipating components or in energy-storing components. This is mainly because the faults were described based on their influences on the system energies. On the other side, Marais’ [16], Uren et al. [53], van Graan [17], Naser [18], and Uren et al. [19] characterised their various systems in terms of energy quantities (in some instances only exergy). The energy properties are then packaged in such a way that the physical process structure information is retained. Consequently, the hybridisation focus on data abstraction in terms of energy characteristics, rather than just combining different techniques [12]. Marais’ [16] work does well in showing the potential of an exergy-based threshold approach when applied to an Autothermal Reformer (ATR). The ATR streams were expressed in terms of physical and chemical exergy, whereafter a threshold function was applied to the data. The approach could successfully detect, and to some extent, isolate the considered faults. Notably, the study shows that systems influenced by chemical variation

would benefit the most from the information encompassed within the chemical exergy. Unfortunately, the approach was not applied to a larger-scale process. In the article published by Uren et al. [53], a counter-flow heat exchanger was characterised and visualised in terms of energy flows. Residuals, in the form of vectors, were then obtained by computing the difference between the normal and faulty conditions. A qualitative error code was lastly generated by testing whether the residuals were above or below a certain threshold value. The successful FDI results demonstrate the abundant information contained within the energy descriptions of a system. Using a similar system, van Graan [17] characterised a counter-flow heat exchanger in terms of exergy and energy flows. The data are then packaged within a linear attributed graph. The graph is used as a framework to store energy information as well as structural connectivity of the system. Using graph theoretical fundamentals, the normal operating condition (NOC) and faulty conditions were visualised in terms of their eigenvalues. Finally, FDI was accomplished by using rule-based fuzzy logic. Expanding on the attributed graph approach, Nesar [18] applied two different graph-based approaches to a Brayton cycle. Both methods made use of exergy and energy flows packaged within an attributed graph. The first approach generated residual-based fault signatures by subtracting considered graph matrices from one another. Excellent FDI results were achieved, with only minor issues regarding multiple simultaneous faults. The second approach made use of eigenvalues and eigenvectors (eigendecomposition) in order to create fault signatures. Although reasonable performance was achieved, the eigendecomposition approach was not as effective as the residual-based approach. In the article published by Uren et al. [19], an energy-based attributed graph matching approach in a heated two-tank system was investigated. The system's exergy and energy flows, for normal and fault conditions, are encapsulated within node signature matrices. Each faulty condition matrix is then compared to the normal matrix to obtain a cost matrix. The eigenvalues of these cost matrices are then assessed qualitatively to obtain a unique fault signature. While favourable fault detection was achieved, the improvement of the isolation capabilities was highlighted as future work.

In light of the foregoing reviewed literature, it is evident that energy is not only a unifying parameter across physical domains but also reduces the data dimensionality. Additionally, energy lends itself to describing various aspects of a system. This includes, but is not limited to, the energy stored, energy dissipated, energy supplied, the system's efficiency and the usefulness of available energy. Hence, energy descriptions could significantly contribute to FDI [12, 52, 54].

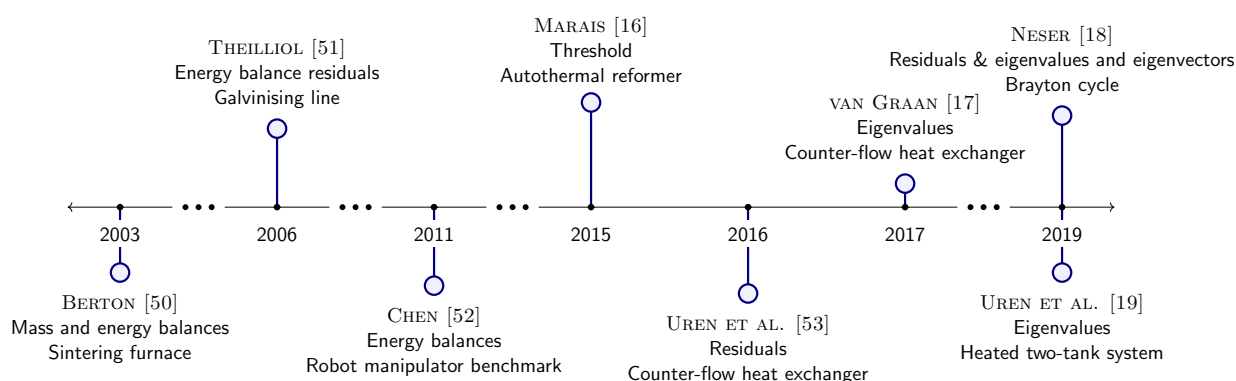


Figure 2.5: Time line of surveyed energy-based FDI approaches

2.4 ADVANTAGES AND SHORTFALLS

As mentioned, all FDD approaches have some advantages and drawbacks. In order to better understand and assess the approaches, the subsequent sections will summarise the strengths and shortfalls seen in literature. As with most things, these attributes are not always applicable to every technique and/or system. The features are given under the relevant FDD category with quantitative model-based methods summarised in Table 2.1, qualitative model-based methods in Table 2.2, data-driven methods in Table 2.3, and hybrid approaches in Table 2.4.

Table 2.1: Quantitative model-based methods advantages and shortfalls

Advantages
• Models are based on physical and mathematical laws and properties [2, 13, 41]. • When adequately constructed, they generate the most accurate output estimations [41]. • Both normal and faulty conditions can be modelled based on first principles, therefore faulty conditions are easily distinguished from normal operating conditions [41]. • These methods can describe dynamic/transient behaviour of the system [13, 41].
Shortfalls
• The mathematics and laws behind the system can be very complex, therefore: - requiring a great deal of effort to develop the model; if at all possible [2, 41]. - rendering the approach computationally intensive [2, 41]. • Sometimes, the approach requires many inputs to describe the system, some of which might not be easily obtainable [41].

Table 2.2: Qualitative model-based methods advantages and shortfalls

Advantages
• Conclusions can be drawn about a system without making use of explicit laws, expert knowledge or numerical values [41]. • Partial conclusions can be drawn about a system from incomplete or uncertain knowledge [31, 41]. • The methods are easy to construct and employ [41]. • The reasoning of these methods is transparent, providing excellent explanation facility because of the cause-effect characteristics [41].
Shortfalls
• The methods applied are tailored to a specific system or process [41]. • It is not easy to ensure that all the described rules are always applicable or complete, especially when the system is complex [41]. • These methods are prone to the generation of spurious solutions [31]. • As rules are modified, included and/or extended to encase new or special conditions, some of the simplicity is lost [41].

Table 2.3: Data-driven methods advantages and shortfalls

Advantages
• They are well-suited in applications where large datasets are cheaply and readily available [41]. • Data-driven approaches proficiently handle problems with no theoretical descriptions of behaviour or its performance [41]. • Development of data-driven techniques requires almost no understanding of the underlying physics of the system [41]. • Most data-driven techniques' computational requirements are minimal [41]. • Some data-driven techniques allow for dimensionality reduction [13].
Shortfalls
• Most data-driven models cannot extrapolate beyond the training data ranges [41]. • Large sets of real plant data are required to train a representative system; usually, these datasets do not include many labelled faulty data and most certainly not the entire possible range of abnormal occurrences [41]. • In order to guarantee robust estimates, the measurement errors of data should be minimal [41]. • The models are system-specific and can seldom be used on other systems [41]. • Some of the data-driven approaches perform well in detecting and isolating faults but have limited success in identification of faults [26].

Table 2.4: Hybrid methods advantages and shortfalls

Advantages
• A hybrid approach would be able to increase the performance of reliability and handling of uncertainties [26]. • Able to utilise and combine the array of different information and data that is available [26]. • Might encapsulate important information of the connectivity/causality of the system [13]. • A hybrid approach would benefit from the progress made in all the various FDD fields [13].
Shortfalls
• The shortfalls of combining various techniques might not all be evident until after development, employment and evaluation. • Based on the encompassed techniques, the hybrid approach might still be system-specific.

2.5 PERFORMANCE CRITERIA

In order to evaluate a proposed FDD system's performance, especially in comparison with another, there needs to be a consensus on what characteristics to inspect, and in which manner. A few propositions are seen in literature, the most popular discussed in the subsequent sections.

2.5.1 PATEL AND KAMRANI [55]

One of the first compilation of FDD system specifications was documented by [55]. The authors - specifically working with expert systems - deemed the specifications tabulated in Table 2.5 a necessity. The proposed properties seem to describe most of the essential functions expected of an FDD system. Some specifications, however, such as "*suggest improvements on the design for maintenance*", fall outside of the modern FDD scope. Although good descriptions of the criteria are given, exact details on how to evaluate and/or quantify

them are somewhat lacking.

Table 2.5: Summary of Patel and Kamrani’s assessment criteria [55]

Level of Performance
• Accurate fault identification • No fault alarm when normal • Degree of confidence (Degree of Confidence) for diagnosis • Possible conclusions ranked by Degree of Confidence • Handle insufficient data and uncertainties
Adaptability
• Diagnose electric, electronic, and mechanical failures • System should easily adapt to changes in a mechanical system • Allow for easy addition, deletion, and modification
Other features
• System should verify it’s sensors’ accuracy (on-line sensing) • Explanation facilities on how diagnosis was reached • Supply operators with applicable recommendations • Simplistic and accessible user interface
Future expectations
• System should be able to plan and control maintenance operations • Suggest improvements on the design for maintenance

2.5.2 VENKATASUBRAMANIAN ET AL. [2]

The ten desirable characteristics an FDD system should demonstrate, as set out by [2], is shown in Table 2.6. The criteria, albeit very comprehensive, once again only assess the qualitative properties of the FDD system. It does not describe the precise evaluation of detection rates or accuracies; these being additional measures that might better illustrate overall performance. Nonetheless, literature would suggest that this list of criteria is well-known and widely used.

2.5.3 REDDY [56]

Another distinguished researcher that developed a generic methodology for assessing FDD systems is T.A. Reddy [56]. The methodology looks at four main criteria categories; (1) site-specific criteria, (2) performance criteria, (3) cost criteria and (4) testing and benefit analysis. The proposed framework was developed explicitly for the heating, ventilating, air-conditioning, and refrigerating (HVAC&R) field, but some of the aspects could be adopted universally. Table 2.7 tabulates the elements of the proposed procedure. The most noteworthy element of the methodology would be that both qualitative and quantitative metrics are evaluated. The author remarks on the fact that the proposed methodology would only be applicable to site-specific HVAC&R systems; therefore a subset of the criteria should be considered if an unrelated and independent FDD system is surveyed. An important metric, highlighted in the quantitative performance section, is the detection and diagnosis outcomes. The fault detection is measured in terms of false positives (FP), false negatives (FN), true positives (TP), true negatives (TN), and no detections (ND). With the diagnosis being either correct diagnosis (CD), misdiagnosis (MD), and no diagnosis (ND_G). A graphical representation of these states are given in Figure 2.6. Each of the terms is briefly discussed, with Figure 2.7

detailing the detection concepts visually.

Table 2.6: Desirable characteristics of an FDD system as proposed by Venkatasubramanian [2]

Quick detection and diagnosis
Quickly detect and diagnose faults
Isolability
Distinguish between different faults
Robustness
Robust to noise and uncertainties
Novelty identifiability
Is the fault known or unknown (novel)
Classification error estimate
Strengthen user's confidence in system's reliability by providing estimates on classification errors
Adaptability
System should be adaptable to changes
Explanation facility
System should justify why certain recommendations were made but also why others weren't proposed
Modelling requirements
The modelling effort should be as minimal as possible
Storage and computational requirements
Dependent on the types of algorithms and implementations a balance between storage and computational requirements should be kept
Multiple fault identifiability
Able to identify multiple faults occurring simultaneously

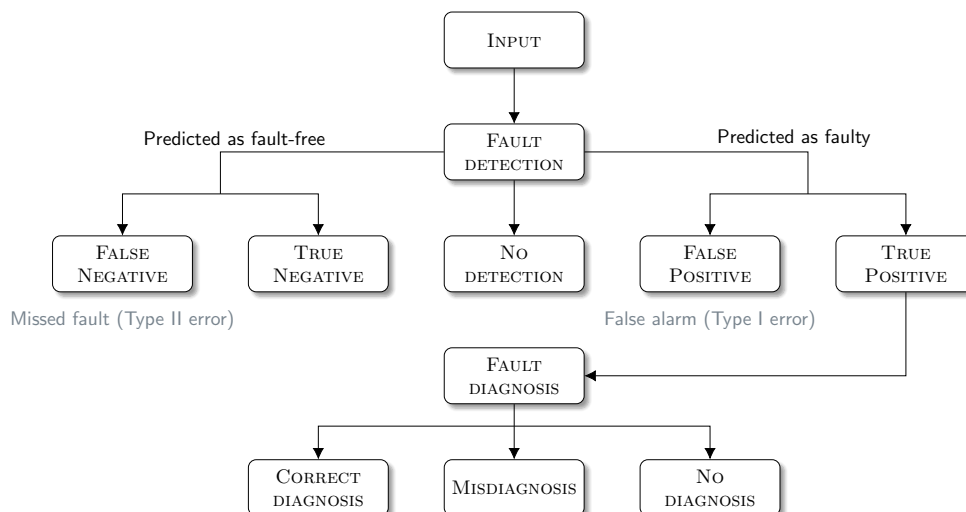


Figure 2.6: Reddy's evaluation procedure for detecting and diagnosing faults [56]

Table 2.7: A summary of Reddy's FDD assessment methodology [56]

Site-specific
• Consider the type of equipment or system • List, in order of importance, plausible faults • List sensors available and accuracy thereof • List already available automated systems • Describe range and frequency of different operating states • Detail annual costs for: - operating energy - maintenance - operator labour with false positive tasks
Performance
Qualitative
• Simplicity in terms of: - understanding the system - utilising the system • Accurate fault identification • Rank possible faults if not uniquely identifiable • Automatically adapt/learn to improve sensitivity and robustness • Handle simultaneous faults
Quantitative
• Low false positives and false negatives • Sensitivity • Rapid identification of abrupt faults
Cost
Qualitative
• Ease of: - integration with existing systems - modification and flexibility in different operating conditions - transportability - routine maintenance - calibrations
Quantitative
• Cost of: - initial FDD system - additional sensors (if required) - implementation and/or commissioning - training operators to used FDD system - delayed benefit of using FDD system (while still training operators) • Operator cost of: - maintenance/repairs - false positive call-outs • Savings due to: - reduced energy use - reduced maintenance costs
Testing sequence & cost benefit analysis
• Test or emulate different faults and operating conditions prior to selling of FDD

2.5.3.1 REDDY'S FAULT DETECTION METRICS

FALSE NEGATIVE (FN)

FN also called a missed fault, or type II error is when the FDD system reports a fault-free state but the true

condition is that a fault exists.

TRUE NEGATIVE (TN)

TN is when the FDD system reports a fault-free state, and the true condition is fault-free.

NO DETECTION (ND)

ND is when the FDD system cannot be applied and/or gives no response.

FALSE POSITIVE (FP)

FP also called a false alarm or type I error is when the FDD system reports a fault but the true condition is fault-free.

TRUE POSITIVE (TP)

TP is when the FDD system reports a fault, and the true condition is that a fault exists.

		TRUE CONDITION	
		Fault-free	Fault
DETECTION INDICATION	Fault-free	True negative TN	False negative FN
	Fault	False positive FP	True positive TP

Figure 2.7: Confusion matrix of the FDD system outputs versus the true conditions

2.5.3.2 REDDY'S FAULT DIAGNOSIS METRICS

CORRECT DIAGNOSIS (CD)

CD is when the FDD system reports a fault type that matches the true condition fault type.

MISDIAGNOSIS (MD)

MD is when the FDD system reports a fault type that does not matches the true condition fault type.

NO DIAGNOSIS (ND_G)

ND_G is when the FDD system cannot provide a diagnosis output.

2.5.4 KURTOGLU ET AL. [57]

In the paper of Kurtoglu et al. [57] the proposed performance metrics are divided into two main categories, namely the detection metrics and diagnosis metrics. Furthermore, a distinction is made between temporal and static metrics. The temporal metrics regard the FDD system's response to a time-varying signal, whereas the static metrics are seen as being independent of time. With this assessment approach, both qualitative and quantitative properties are evaluated. The temporal metrics, however, would become less relevant in systems with large time constants. The details of the metrics are given in Table 2.8

Table 2.8: Assessment metrics as detailed by Kurtoglu et al. [57]

Detection metrics
• Detection response time • Detection false positive rate • Detection false negative rate • Detection rate • Detection accuracy • Sensitivity • Stability
Isolation metrics
• Isolation response time • Time to isolate • Time to estimate • Accuracy • Resolution • Stability

2.6 APPLICATIONS

When surveying the literature relevant to process monitoring, it quickly becomes apparent that many different systems and techniques have been considered over the years. This includes applications in aerospace, automotive, chemical, electrical, pharmaceutical, HVAC&R, mechanical, and nuclear fields of study. Severson et al. [26] state that by March 2015, over 34 000 related publications were produced since the 1970s. It is thus impossible to acquire and study all of them. With Sasol¹ being the financial benefactor, a study within the chemical domain would be advantageous to the company and research community alike. The author started by obtaining the most recent and prevalent review articles, which included the work done by [2, 13, 26, 31, 32, 58, 59]. Some of the articles that these reviewers deemed important and which were distinctly within the chemical field, were collected and evaluated. Figure 2.8 summarises the findings, showing some of the popular chemical systems and the FDD techniques associated with them. This evaluation is by no means all-inclusive but serves as a foundation in determining which system would best further the knowledge base.

One of the most widely used systems is the Tennessee Eastman Process (TEP) [46,47,103–118]. The process was constructed by the Eastman Chemical Company to provide a practical source of industrial process data for evaluating different control and monitoring techniques. The developed simulation is based on a realistic chemical process; existing of a reactor, condenser, compressor, separator and stripper. The process contains eight components, labelled A-H, with the particulars of these components, kinetics and operating points obscured for proprietary purposes [119]. There exist a few simulation variations of the TEP, a popular one being the datasets that were published by Russell et al. [119]. In 2008 Lin et al. [120] developed a model in the commercial process simulator Aspen Plus[®]. Nonetheless, the authors do not substantiate the reactions, and proposed components (A-H) used, nor give any specifications of the final flowsheet. This makes the model unduplicatable. Another model, developed in Modelica, was published in 2018 by [121]. The aim

¹Sasol is an international integrated chemical and energy company based in South Africa.

was to provide an object-oriented model that was freely available to use as a benchmark. By not knowing all the specifics of the TEP, it would be a complex effort to reproduce the model within a commercial process simulator such as Aspen HYSYS[®][122].

The other systems such as Continuous stirred-tank reactors (CSTR) [79–88], reformers [16, 78] and heat exchangers [17, 89–91] work well in the early developmental stages of a proposed FDD system. These systems are usually sufficient for testing the intended concepts but lack the necessary complexity larger scaled systems provide.

2.7 CONCLUSION

When considering the impact unreliable operations of a system can have on safety and financial aspects, the benefit of automated process monitoring becomes clear. For the past 50 years, researchers have generated an array of model-based and data-driven techniques to accomplish FDD tasks. Recently, substantial advances were made in developing hybrid techniques, which aim at addressing the drawbacks single approaches indicated. Most researchers are of the opinion that hybridisation is the way forward [13, 26]. Of specific interest is the energy-based hybridisation; the work done thus far, demonstrating compelling results [17, 18, 53, 123, 124]. The most notable being the encapsulation of process structure information and energy characterisation which serves as a universal description of the system. As such, these energy-based FDI approaches should be evaluated within a more complex system. Literature would suggest the use of the benchmark Tennessee Eastman Process (TEP). However, with the exact specifics thereof obfuscated, it would be impractical to attempt to reproduce a model within a commercial simulation environment. Seeing as the Autothermal Reformer (ATR) Marais [16] worked on is the first process unit within a gas-to-liquids (GTL) process; it would be pragmatic to develop and use a complete GTL process. The GTL process would be sufficiently complex, and enough process information is available to build a representative model with. Chapter 3 will, therefore, look at the development of a GTL process to use as a base for this study.

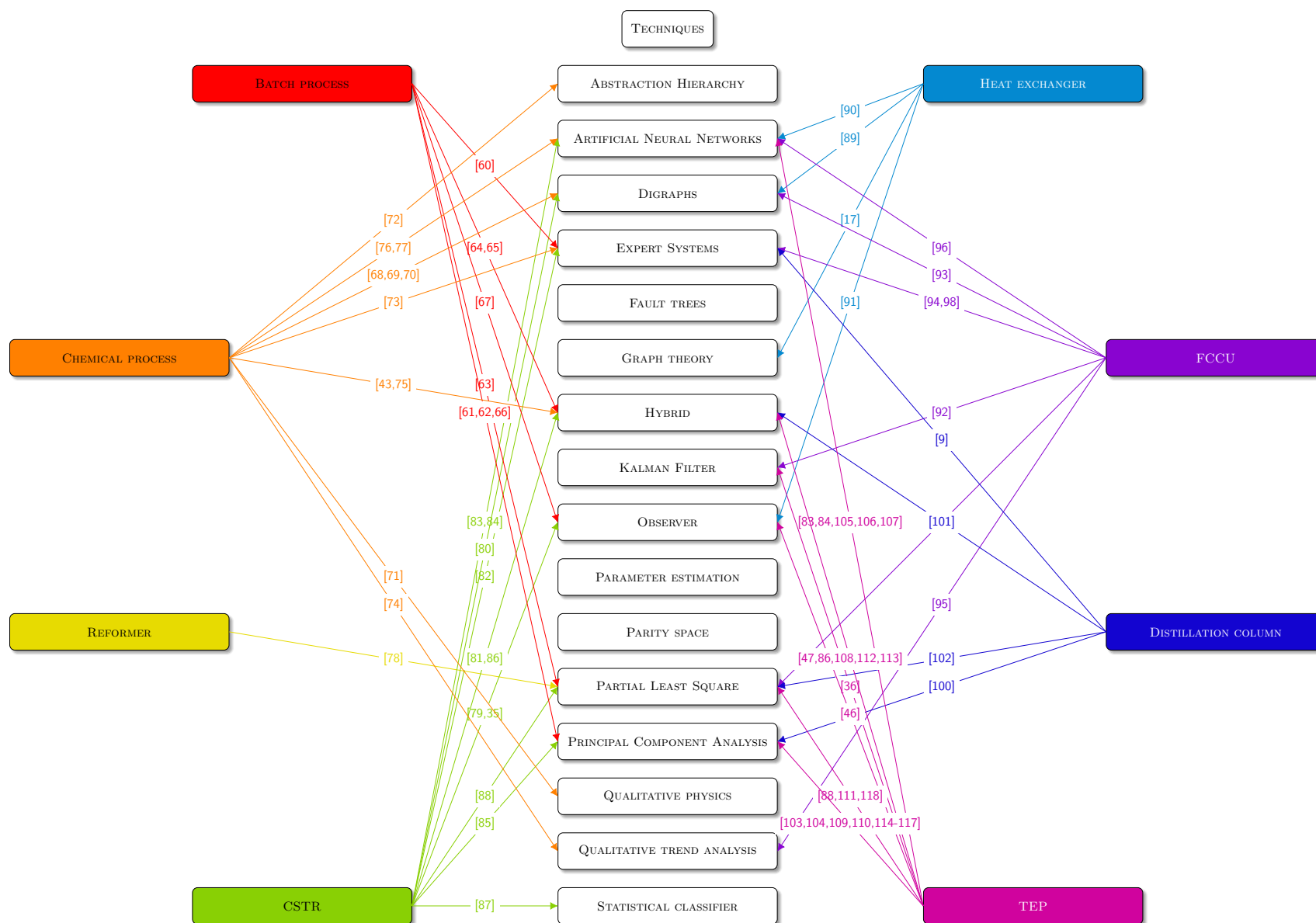


Figure 2.8: Different chemical systems and the prominent FDD techniques found in literature

CHAPTER 3

GAS-TO-LIQUIDS MODEL

3.1 INTRODUCTION

The previous chapter surveyed the prominent FDD literature, showing the various existing approaches and applications. The GTL process was chosen as a suitable larger-scale system to use as a basis for this study. As such, this chapter is dedicated to the GTL model development and related discussions. The chapter begins by looking at synthetic fuel and its history. Next, a general process overview and terminology are given. The GTL's overall process and comprising sections are discussed in slightly more detail. The focus is then shifted to the development of the simulation model of the GTL process. The modelling assumptions are highlighted and substantiated before moving on to the particulars of the modelling effort. The model validation is also reviewed. As with any FDI investigation, faulty process conditions are required. The central units of a GTL process are identified after which common failures and causes are examined. Based on these findings, fault conditions are formulated and detailed.

3.2 SYNTHETIC FUEL

3.2.1 HISTORICAL BACKGROUND

By the 1920s petroleum had become an integral part of the economies of industrialised countries. The major transition from solid to liquid fuels was seen mainly due to the high energy potential contained within petroleum as opposed to that of coal or wood. Furthermore, the advances made in the automotive, aircraft and maritime industries also played a considerable role. Nations such as Britain, Canada, France, Germany, Japan, and Italy had mostly imported naturally-occurring petroleum, as they had limited to no access to domestic petroleum [125]. During 1914 - 1919 the British established a naval blockade, referred to as the Blockade of Germany, which strived to cut off all commerce to Germany, Austria-Hungary, the Ottoman Empire, and Bulgaria. This blockade had crippling effects as no petroleum could be imported during this time. Even after the blockade was lifted, their economy did not allow for the acquisition of foreign fuels [24]. This, amongst others, motivated Germany to endeavour to produce liquid petroleum from coal. Two specific processes, contrived quite some time before, laid the cornerstone of the successful synthesizing of

Table 3.1: Companies and their plants, years operational, Barrels per day capacity, type of feedstock, and location [24, 127–131]

Company	Plant	Year	Capacity (Barrels per day)	Feedstock	Location
ExxonMobil	SynFuels 600 KTA	1985-1997	14 500	Methanol	New Zealand
	/	2009	2 500	Coal	China
	/	2015	/	Methanol	China
PetroSA	Mossgas	1992	25 000	Natural gas	Mossel Bay
Rentech	/	2000	/		Colorado
Sasol	Sasol I	1955-1993	2 500	Coal	Sasolburg
		1993-2004	/	Coal	
		2004-present	5000	Natural gas	
	Sasol II	1980	85 000	Coal	Secunda
	Sasol III	1982	85 000	Coal	Secunda
Qatar Petroleum & Sasol	Oryx GTL	2007	34 000	Natural gas	Qatar
Qatar Petroleum & Shell	Pearl GTL	2011	140 000	Natural gas	Qatar
Shell	Bintulu	1993	14 700	Natural gas	Bintulu
Shenhua Group	Shenhua DCL	2010	24 000	Coal	Inner Mongolia
Syntroleum	Dynamic Fuels	2010	2 500	Animal feedstocks	Louisiana

3.2.2 GENERAL PROCESS

In order to obtain synthetic fuels, feedstock with high carbon and hydrogen content is converted via an appropriate chemical process. There exist a number of different routes that can achieve this; the main approaches being *direct conversion* and *indirect conversion*. *Direct conversion* transforms a feedstock directly to fuel without the need to produce any medial gasses. One example of such a process is when methane (CH_4) is transformed directly into C_2 -hydrocarbons or methanol (CH_3OH). The general consensus, however, deems *direct conversion* technology commercially infeasible; the only commercial plant currently operational being the Shenhua Group direct coal liquefaction (DCL) demonstration [127]. The more popular route, *indirect conversion*, is the gasification of a carbonaceous feedstock to obtain synthesis gas (syngas) which is then further converted either via Fischer-Tropsch (FT) or the Mobil process, to produce synthetic fuels (also called syncrude). The syncrude is then upgraded and refined to produce the desired fuels and chemicals [132, 133]. This process is alternatively known as the feed-to-liquid (XTL) conversion. Figure 3.2 graphically shows the typical process flow and comprising sections of an XTL process; the feedstock X either being $\text{C} \Rightarrow$ coal, $\text{G} \Rightarrow$ natural gas, $\text{B} \Rightarrow$ biomass, or $\text{W} \Rightarrow$ organic waste. The specific conversion this study is considering, is the gas-to-liquids (GTL) route. The subsequent section will discuss the relevant technology and applications thereof.

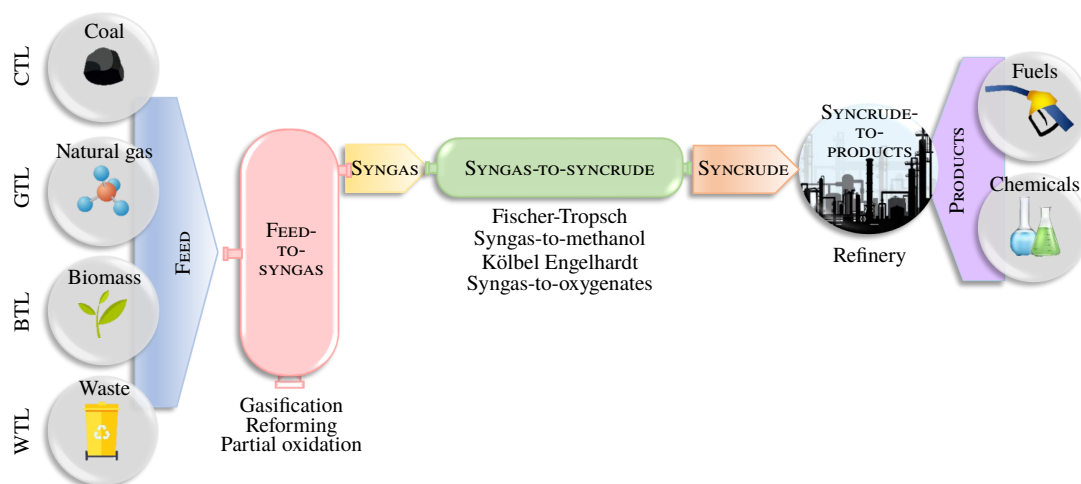


Figure 3.2: Indirect conversion of carbonaceous feedstock to synthetic fuels and chemicals (Adapted [132, 133])

3.3 GAS-TO-LIQUIDS

A gas-to-liquids (GTL) process, which is categorised as an *indirect conversion*, usually consists of three main processing sections as shown in Figure 3.3. Gaseous feedstock, such as natural gas, is firstly converted to synthesis gas. This synthesis gas is then transformed by making use of Fischer-Tropsch (FT) reaction to obtain liquid hydrocarbons of various chain-lengths. Finally, cracking and hydro-processing are utilised to upgrade the products to desired specifications [22, 24]. For each of these processing sections, there exist many different architectures and approaches. The subsequent subsection will discuss the various sections' constituents and relevant details.

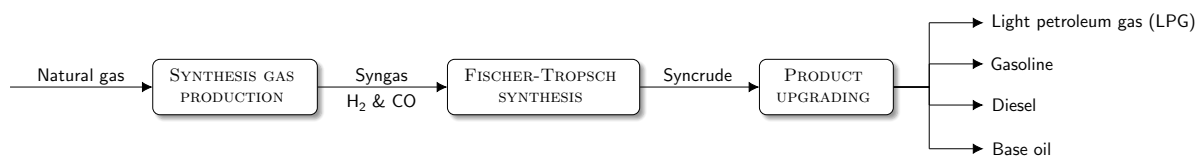


Figure 3.3: The three major sections of a GTL process (Adapted [22])

3.3.1 SYNTHESIS GAS PRODUCTION

As mentioned, the synthesis gas production section converts natural gas to synthesis gas, also known as syngas. Syngas is a mixture of hydrogen (H_2) and carbon monoxide (CO) in a specific ratio to one another. Chemical reforming is generally used to produce syngas. This processing section is known to be the most expensive step, contributing to over half of the total capital cost. As a result, a wide array of technologies were researched and developed over the years [24, 134]. The syngas production usually comprises of a pre-reformer, reformer and cleaning/conditioning sections.

3.3.1.1 PRE-REFORMING

Natural gas does not consist solely of methane (CH_4); traces of ethane (C_2H_6), propane (C_3H_8), butane (C_4H_{10}), nitrogen (N_2), and sulphurous compounds can also be found [16, 24]. To avoid cracking of heavier hydrocarbons within the syngas reforming section, a pre-reformer is employed [23]. A secondary advantage of a pre-reformer is the chemisorption of undesired sulphur within the feedstock [24].

3.3.1.2 REFORMING

For the main syngas production there exist quite a few reforming routes, each having distinct advantages and disadvantages. These reforming pathways are usually classified as being either catalytic or non-catalytic, the categorisation based on the utilisation of a catalyst. The most prevalent technologies seen throughout literature is Steam methane reforming (SMR), Autothermal reforming (ATR), Heat exchange reforming (HER), Non-catalytic partial oxidation (POX), Catalytic partial oxidation (CPO) [24, 133, 135]. Many researchers have - in great detail - discussed and evaluated the various technology. Therefore, Table 3.2 gives an overview of only the most important aspects.

3.3.1.3 SYNGAS CLEANING AND CONDITIONING

It is well-known that sulphur and nitrogen-containing compounds (chlorides and bromides) degrade Fischer-Tropsch (FT) reactor catalysts. To mitigate these effects, the syngas is usually cleaned as an intermediate step before being fed to the FT reactor. This processing step also allows for conditioning of the syngas, if necessary, by adjusting its composition, temperature, etc. [24, 132].

Table 3.2: The various syngas production reforming technologies (Adapted [24, 133, 135])

Technology	Reactants	Catalyst	Reaction type	Operating		Syngas ratio (H ₂ /CO)	Advantages	Disadvantages
				Temperature [°C]	Pressure [kPa]			
SMR	→ Methane → Steam	Nickel-based	Endothermic	800 - 900	2026.50 - 3039.75	≈ 3	→ Most extensive industrial experience → No oxygen required → Lowest process temperature requirement → Best H ₂ /CO ratio for hydrogen production applications	→ H ₂ /CO ratio higher than required → Highest CO ₂ emissions
POX	→ Methane → Oxygen	Non-catalytic	Exothermic	1200 - 1400	\	< 2	→ Feedstock desulfurisation not required → Absence of catalyst permits carbon formation and, therefore, operation without steam, significantly lowering syngas CO ₂ content → Low methane slip → Low natural H ₂ /CO ratio is an advantage for applications requiring ratio < 2	→ Low natural H ₂ /CO ratio is a disadvantage for applications requiring a ratio < 2 → Very high process temperatures → Usually requires oxygen → High-temperature heat recovery and soot formation/handling adds process complexity → Syngas methane content is inherently low and not easily modified to meet downstream processing requirements
CPO	→ Methane → Oxygen	→ Platinum → Palladium → Rhodium → Iridium	Exothermic	700 - 1000	\	\	→ Lower temperatures than ATR → Lower oxygen consumption	→ Cost of the catalyst (usually a noble metal, in particular Rh)
ATR	→ Methane → Steam → Oxygen	Nickel-based	Endothermic-Exothermic	950 - 1050	2000 - 4000	≈ 2	→ Natural H ₂ /CO ratio is often favourable → Lower process temperature requirement than POX → Low methane slip → Syngas methane content can be tailored by adjusting reformer outlet temperatures	→ Usually requires oxygen
HER	→ Methane → Steam → Oxygen	Configuration specific	Endothermic-Exothermic	\	\	\	→ Compact overall size → Application flexibility offers additional options for providing incremental capacity	→ Limited commercial experience → In some configurations, must be used in tandem with another syngas generation technology

3.3.2 FISCHER-TROPSCH SYNTHESIS

3.3.2.1 THE PROCESS

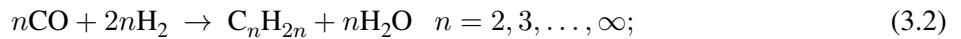
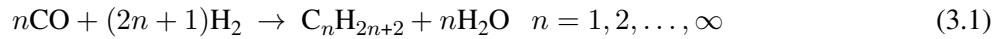
Fischer-Tropsch synthesis (FT) is the catalytic transformation of syngas to hydrocarbons of various chain-lengths. Similar to the syngas production, different pathways of converting syngas to syncrude exist. Typically FT technology is described by the reactor type, catalyst and operating conditions. The aim is to determine a combination of these aspects in order to produce the desired syncrude composition [24, 132]. The syngas composition is another aspect that influences the syncrude obtained. Since the FT reaction is highly exothermic and the product selectivity temperature-dependent, the reactor types that can be utilised, are limited to the heat management properties thereof. The catalyst selection also plays a major role in the reactor type as the catalyst deactivation rate and replacement strategy determine the reactor's suitability. FT reactors that are seen throughout literature (and commercially) are fixed beds (multi-tubular and microchannel), fluidized beds (fixed and circulating), and slurry beds. The operating conditions greatly affect the FT catalysis, which directly influence desorption, hydrogenation, and the chain growth probability. For FT synthesis the operating temperature is used to classify the technology. The categories being low-temperature Fischer-Tropsch (LTFT), medium-temperature Fischer-Tropsch (MTFT), and high-temperature Fischer-Tropsch (HTFT) synthesis [133]. The temperature range for every classification is summarised in Table 3.3. Specialised and detailed discussions are given in the works of [126, 132, 133, 136]. Thus, for the purpose of this study, a summary of the popular commercial FT technologies is tabulated in Table 3.4.

Table 3.3: FT synthesis classification temperatures

Synthesis	LTFT	MTFT	HTFT
Temperature range [°C]	< 250	≈ 270	> 320

3.3.2.2 ANDERSON-SCHULZ-FLORY DISTRIBUTION

The polymerisation reactions taking place within the FT reactor are hydrogenation of CO that forms n -paraffins, 1-olefins. Assessing the respective generic equations:



it is evident that an infinite number of reactions exist. These reactions also describe little in terms of the product distribution. A widespread assumption seen in literature, suggests that the ratio between two consecutive reaction rates is quantified by a constant called the growth factor (α). This assumption is applied in order to achieve a model that would be finite, but would also consider all reactions and components [137]. To model the product distribution the Anderson-Schulz-Flory distribution model is used. The model, shown in (3.3), describes the distribution of the product weight fractions (w_n) of considered C_n , as a function of the carbon number (n).

Table 3.4: Different FT technologies seen commercially over the years [24, 127–130]

Year	Company/process name	Synthesis category	Reactor type	Catalyst
1936	German normal-pressure	LTFT	Fixed bed	Cobalt
1937	German medium-pressure	LTFT	Fixed bed	Cobalt
1951	Hydrocol	HTFT	Fixed fluidised bed	Iron
1955	Arbeitsgemeinschaft Ruhrchemie-Lurgi	LTFT	Fixed bed	Iron
1955	Kellogg Synthol	HTFT	Circulating fluidised bed	Iron
1955	Sasol I	HTFT	Circulating fluidised bed and Multi-tubular fixed bed	Iron
1980	Sasol (II) Synthol	HTFT	Circulating fluidised bed (SAS)	Iron
1982	Sasol III	HTFT	Circulating fluidised bed (SAS)	Iron
1985	SynFuels 600 KTA	-	Fixed bed	ZSM-5
1992	Mossgas	HTFT	Circulating fluidised bed	Iron
1993	Sasol Advanced Synthol (SAS)	HTFT	Fixed fluidised bed	Iron
1993	Shell middle distillate synthesis	LTFT	Fixed bed	Cobalt
1995	Iron Sasol slurry bed process	LTFT	Slurry bubble column	Iron
2000	Rentech	-	Slurry bubble column	Iron
2004	Sasol I	LTFT	-	Iron
2005	Statoil cobalt slurry bubble column process	LTFT	Slurry bubble column	Cobalt
2007	Onyx GTL	LTFT	Slurry bubble column	Cobalt
2008	High-temperature slurry FT process	MTFT	Slurry bubble column	Iron
2011	Pearl GTL	LTFT	Multi-tubular fixed bed	Cobalt
2015	ExxonMobil		Fluidised bed	ZSM-5

The chain growth probability factor (α) depicts the probability of a sequential propagation step and is dependent on the type of catalyst utilised.

$$w_n = n(1 - \alpha)^2 \alpha^{n-1} \quad (3.3)$$

α -values of 0.95 or higher are normally seen in LTFT processes [24]. A diagrammatic representation of the model is shown in Figure 3.4. It should be noted that the model describes the ideal product distribution and has shown to deviate slightly from experimental findings. Alternative methods of determining appropriate α -values are discussed in [23].

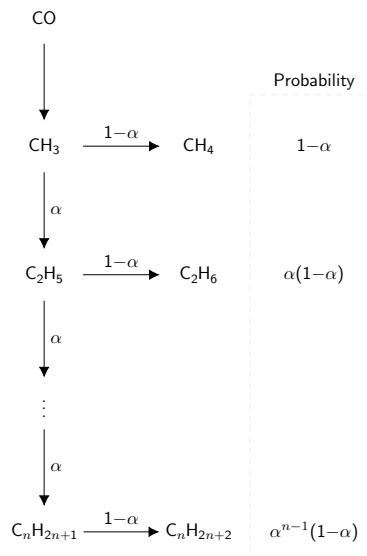


Figure 3.4: Probability of chain growth to subsequent hydrocarbons in FT reactions [23, 24]

3.3.3 PRODUCT UPGRADING

The final step in a GTL process is the upgrading or refining of the obtained syncrude. According to [132], three general levels of syncrude-to-product conversions exist. Table 3.5 summarises these conversions; the main difference being whether the products obtained are intermediate or final. It is interesting to note that commercial FT facilities incorporate at least a form of partial refining. As this processing section is not included in this study's scope of work, no further descriptions are given. The interested reader is referred to the comprehensive insights of [132].

3.4 DEVELOPED GTL MODEL

The following section is used to detail the various aspects of the developed GTL model. It starts off by stating the basis of the model such as the simulation software and modelling assumptions. Additionally, an overview of the modelled process and its general layout are given. The precise detail regarding the reformer and FT reactor simulation are also described. The model validation is documented in Section 3.4.6 and the section is concluded with the specifics of the recycling section.

Table 3.5: Syncrude-to-product conversions and product details

Conversion	Details	Type of product
Upgrading	All products still need to be refined in order to obtain final products.	Intermediate
Partial refining	Some products are to be refined while others are blended to produce final products.	Intermediate and final
Stand-alone refining	All obtained products are final products.	Final

3.4.1 SIMULATION SOFTWARE

When surveying relevant literature, the popular modelling software academia use to simulate processes are Honeywell's UniSim[®], Aspen Plus[®], and Aspen HYSYS[®] [16,20–24,138]. All of these commercial process simulators:

- are capable of mathematically simulating complex chemical processes, making it possible to model a single process unit, a large-scale chemical plant or a refinery.
- include integrated tools for costing, energy management, and safety analyses.
- were developed for using extensively within the gas and oil field.

Thus, any one of the software programs would allow for the simulation of a representative GTL process. For this study, Aspen HYSYS[®] is used, as the university offers students licensed access to the software.

3.4.2 MODELLING ASSUMPTIONS

As with most modelling endeavours, assumptions are made in order to simplify the task at hand [139]. The subsequent sections detail the various assumptions that were established to develop representative, steady-state GTL model within HYSYS[®].

3.4.2.1 FEEDSTOCKS

As mentioned, natural gas consists primarily of methane. Additionally, it usually contains higher hydrocarbons and other impurities such as sulphur and nitrogen compounds. The specific composition of natural gas can vary not only by region but also over time. To ensure the developed model for this study was kept simplistic enough, no pre-reformer is to be included. This means no cleaning and conditioning of the natural gas was modelled and therefore pure methane (CH₄) was used as feedstock. All other input feedstocks used were also pure.

3.4.2.2 PROCESS FLOW

The modelled GTL process consisted of a reformer for syngas production, a reactor for achieving the Fischer-Tropsch (FT) synthesis and the recycling of unreacted gas to the FT reactor only. To simplify the process

slightly, no pre-reformer was employed as (a) the feedstock requires no cleaning as pure methane is used and (b) no recycling to the reformer was included. The upgrading section was additionally discarded as the complexity of this processing step, illustrated in [138], exceeds the scope of this study.

3.4.2.3 THERMODYNAMIC PACKAGE

When starting a HYSYS[®] model, the user has to choose an appropriate *fluid package* to work with. This package, also referred to as the thermodynamic model, forms the basis for the physical properties of components and mixtures as functions of temperature and pressure. Every package is uniquely suited to certain types of components and operating conditions. If the incorrect package is selected, erroneous simulation results might be obtained [140, 141]. These packages are usually categorised as being either Equation of State (EOS), vapour pressure or empirical. If working with hydrocarbons, as one would when considering a GTL process, an Equation of State (EOS) method would be best suited. HYSYS[®] provides a few different EOS methods, with the Peng-Robinson (PR) hallmarked as the most advanced and an excellent standard for using with hydrocarbons [141]. Based on these facts, the Peng-Robinson thermodynamic package was used throughout the GTL simulation model.

3.4.2.4 NOISE

The well-known types of noise seen in literature are numerical noise and sensor noise. Numerical noise usually comes into effect when working with simulation models, especially if the models make use of differential equations (DEs). This is because DEs are adaptive in nature and might not always converge [142]. Sensor noise is seen as any undesired deviation in a sensor's output without the actual measurand changing. There exist a few different sources of sensor noise as well as methods to counter them [18, 143]. Seeing that this study uses a simulation model, it is assumed that no sensor noise is present within the measurements. The solver variations (numerical noise) noticed within HYSYS[®], however, were regarded small enough to be non-influential.

3.4.3 MODELLED PROCESS AT A GLANCE

As evident from Table 3.2, a few different syngas production pathways exist. When reviewing existing literature, most researchers implement Autothermal reformers (ATRs) [16, 20–24]. ATRs have several advantages amongst which are their economy of scale, smaller footprint, and faster start-up and load transitions [144, 145]. Other authors have also suggested that ATRs show the most promise in terms of GTL processing [146]. Given the suitability of the ATR for use in a GTL process that is fed by natural gas and the advantages it offers at a large-scale for single process streams, this study implements an ATR. De Klerk [132] emphasised the importance of the temperature and composition (ratio of H_2/CO) of the syngas produced. For the specific GTL configuration considered (shown in Figure 3.5), it is expected that the syngas temperature should vary within the range of 1020-1065 °C with $H_2/CO \approx 2.0$. To produce syngas of adequate temperature and composition, the ATR is fed specific ratios of natural gas, steam, and oxygen. It has been shown that oxygen (O_2) greatly affects the syngas temperature, and in some studies, such as [16], a carbon

dioxide (CO₂) stream was included to aid in the control of the syngas composition [138]. The produced syngas is then cleaned (38 °C) by separating out some of the water. The cleaned syngas is then fed at temperatures between 200-240 °C into the Fischer-Tropsch reactor (FTR), categorising it as LTFT synthesis. The considered hydrocarbons included C₂ to C₂₀. C₃₀ was used to represent hydrocarbons C₂₁₋₃₀ which exhibit similar properties. The generation of these hydrocarbons followed an Anderson-Schulz-Flory (ASF) distribution, relating closely to the distributions seen in [22,24]. Usually, unreacted components are recycled to be put through the process again, whilst the liquid products are transferred to the upgrading section.

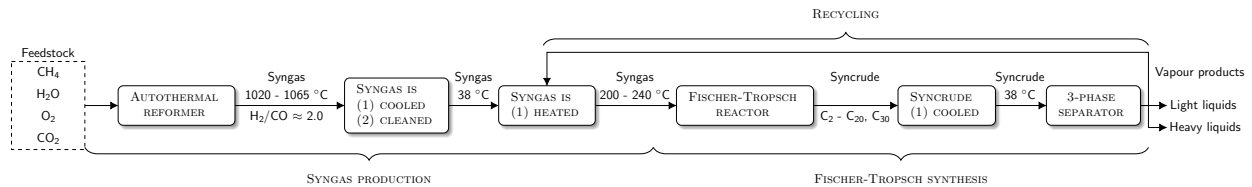


Figure 3.5: An overview of the process flow of the developed GTL process

3.4.4 AUTOTHERMAL REFORMER

To model the ATR, an adiabatic equilibrium reactor was used within HYSYS[®]. According to [23], the three most important equilibrium equations used to describe the ATR reactions are the oxidation of methane (3.4), the steam reforming of methane (3.5), and the water gas shift reaction (3.6):



The ATR and associated conditioning units, extracted from HYSYS[®], is shown in Figure 3.6. The components fed to the ATR were pure methane, steam, oxygen, and carbon dioxide.

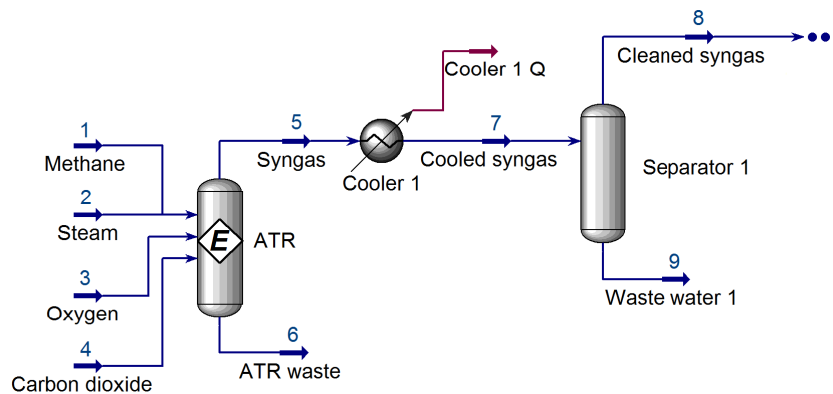


Figure 3.6: HYSYS[®] process flow diagram of syngas production section with relevant stream numbers

With the addition of the carbon dioxide stream, the suggested feed rates of the steam and oxygen would no longer produce syngas of expected temperature or composition. It was anticipated that the feed-ratios would

stay within the same ranges, but to determine the exact feed flow rates, a systematic approach was required. To theoretically determine the new flow rates, the resultant syngas temperature, and syngas composition, Central Composite Rotatable Design (CCRD) was applied. The details and outcomes of this endeavour are documented in Appendix A. After the successful utilisation of CCRD, the final feed ratios are shown in Table 3.6. The corresponding flow rates derived from these ratios are tabulated in Table 3.7. The component temperatures and pressures are also shown here.

Table 3.6: Feed ratios of the ATR components

Ratio	
H ₂ O/CH ₄	0.6625
O ₂ /CH ₄	0.5450
CO ₂ /CH ₄	0.1074

Table 3.7: Syngas production section stream information as simulated in HYSYS®

Stream no	Description	Temperature [°C]	Molar flow [kgmole/h]	Pressure [kPa]
1	Methane	675	8195.0	3000
2	Steam	675	5429.2	3000
3	Oxygen	200	4434.9	3000
4	Carbon dioxide	675	959.4	3000
5	Syngas	1029	30262.6	3000
7	Cooled syngas	38	30262.6	3000
8	Cleaned syngas	38	24452.7	3000

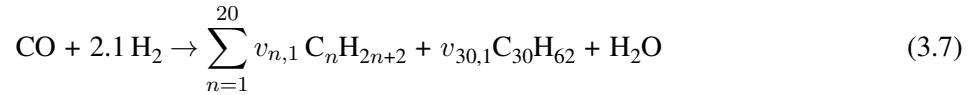
Assessing the syngas stream ([Stream 5](#)), the temperature is seen to be 1029 °C. Furthermore, when evaluating the molar fractions of the syngas, shown in Table 3.8, the composition was found to be $H_2/CO = \frac{0.5035}{0.2396} = 2.105$. Seeing as syngas temperature of ≈ 1030 °C ensures soot-free operation and the ideal composition is ≈ 2 [24, 132], the syngas production was deemed appropriately simulated. To clean and condition the syngas in preparation for the Fischer-Tropsch synthesis section, a cooler ([Cooler 1](#)) was firstly used to cool the syngas down to 38 °C. At this temperature, the steam present is converted to water which can be removed using a separator process unit within HYSYS® ([Separator 1](#)). The waste heat generated by the cooler is not used within the process and is essentially returned to the environment. From a plant design perspective, this is inefficient. However, in doing so, the simulation is kept as simple as possible. Similar arguments hold for the energy streams of compressors and coolers used elsewhere in the process.

Table 3.8: Simulated molar fractions of the main components of the syngas ([Stream 5](#))

Component		Molar fraction
Methane	CH ₄	0.0118
Carbon monoxide	CO	0.2396
Carbon dioxide	CO ₂	0.0512
Hydrogen	H ₂	0.5035
Steam	H ₂ O	0.1940
Oxygen	O ₂	0.0000

3.4.5 FISCHER-TROPSCH REACTOR

The cleaned syngas first needs to be heated before being fed to the Fischer-Tropsch reactor (FTR). Therefore a heater ([Heater 1](#)) is included in the simulation. The feedstock was heated to a temperature of 210 °C [24]. The reactor feed pressure also needed to be lower to 2000 kPa. This was accomplished by setting the *Delta P* parameter within the heater unit. A plug flow reactor (PFR) was used within HYSYS®, as literature suggests that it is representative of a multi-tubular fixed bed (MTFB) reactor. The reactor ([FTR](#)) was implemented with a volume of 1000 m³ and a pressure drop of 60 kPa. Equations (3.7) and (3.8) were modelled as kinetic reactions. Equation (3.7) describes the Fischer-Tropsch reaction (only considering paraffins) where (3.8) represents the inevitable production of methane:



To determine the stoichiometric coefficients of (3.7), the modelling approach proposed by [137] and employed by [24], was used. Similar to these researchers, a constant chain growth factor $\alpha = 0.9$ was utilised. Equation (3.9) is used to calculate all C_n components where $n \leq 20$. The C_{21-30} components are lumped together as a single component ($\text{C}_{30}\text{H}_{62}$) and is calculated using (3.10). The computed coefficients are shown in Table 3.9 and compare well to the values documented in [24].

$$r_{FT} = (1 - \alpha)^2 \alpha^{1-n} \quad \text{for } \text{C}_n \quad \text{where } n = 1, \dots, N \quad (3.9)$$

$$r_{FT} = (1 - \alpha) \alpha^{20} \quad \text{for } \text{C}_n \quad \text{where } n = N + 1, \dots, \infty \quad (3.10)$$

The next step was to specify the rate expressions of the equations. With so many different kinetic mechanisms presented in literature, the most popular approach seems to be that developed by Iglesia et al. [147], given in (3.11) and (3.12)

$$r_{\text{CH}_4} = \frac{k_1 P_{\text{H}_2} P_{\text{CO}}^{0.05}}{1 + K_1 P_{\text{CO}}} \quad (3.11)$$

$$r_{\text{CO}} = \frac{k_2 P_{\text{H}_2}^{0.06} P_{\text{CO}}^{0.65}}{1 + K_1 P_{\text{CO}}}. \quad (3.12)$$

The consensus amongst researchers is to convert these rate of expressions to more universal units [21,23,24]. The values and corresponding units, as used within HYSYS®, are tabulated in Table 3.10. In order to maintain a reactor temperature of ≈ 210 °C, an appropriate *direct Q value* was assigned to the PFR's *energy stream*. An extract of the process flow diagram showing the Fischer-Tropsch section is given in Figure 3.7. The corresponding stream information is tabulated in Table 3.11.

Table 3.9: The stoichiometric coefficients of the CO consumption (Equation (3.7))

n	C_nH_{2n+2}	$v_{n,1}$
1	CH ₄	0.0100
2	C ₂ H ₆	0.0090
3	C ₃ H ₈	0.0081
4	C ₄ H ₁₀	0.0073
5	C ₅ H ₁₂	0.0066
6	C ₆ H ₁₄	0.0059
7	C ₇ H ₁₆	0.0053
8	C ₈ H ₁₈	0.0048
9	C ₉ H ₂₀	0.0043
10	C ₁₀ H ₂₂	0.0039
11	C ₁₁ H ₂₄	0.0035
12	C ₁₂ H ₂₆	0.0031
13	C ₁₃ H ₂₈	0.0028
14	C ₁₄ H ₃₀	0.0025
15	C ₁₅ H ₃₂	0.0023
16	C ₁₆ H ₃₄	0.0021
17	C ₁₇ H ₃₆	0.0019
18	C ₁₈ H ₃₈	0.0017
19	C ₁₉ H ₄₀	0.0015
20	C ₂₀ H ₄₂	0.0014
30	C ₃₀ H ₆₂	0.0122

Table 3.10: The kinetic values and units used for the reactions in HYSYS®

Parameter	Arrhenius Expression		Unit
	A	E	
k_1	8.8×10^{-6}	37326	$\frac{\text{kgmoleCH}_4}{\text{Pa}^{1.05}\text{m}^3\cdot\text{s}}$
K_1	1.1×10^{-12}	-68401.5	Pa^{-1}
k_2	1.6×10^{-5}	37326	$\frac{\text{kgmoleCO}}{\text{Pa}^{1.25}\text{m}^3\cdot\text{s}}$

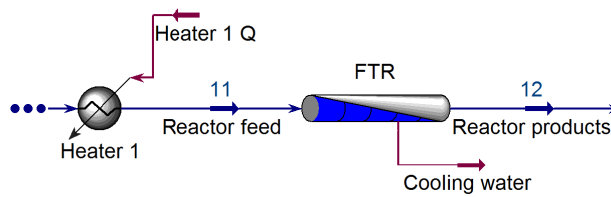


Figure 3.7: HYSYS® process flow diagram of Fischer-Tropsch section

Table 3.11: Fischer-Tropsch synthesis stream information as simulated in HYSYS®

Stream no	Description	Temperature [°C]	Molar flow [kgmole/h]	Pressure [kPa]
11	Reactor feed	210	34310.3	2000
12	Reactor products	213	22290.4	1940

3.4.6 MODEL VALIDATION

To validate whether the model produces the expected products, the product distribution was evaluated. This was done by assessing the weight fractions of the components in [Stream 12](#); before simulating any recycling aspects. The weight fractions (w_n) were firstly divided by their corresponding carbon numbers (n) and then the logarithm of each was calculated. The obtained values are summarised in Table 3.12. The log-values were then plotted against their carbon numbers, as depicted in Figure 3.8. For the ASF distribution, a straight line with slope $\log(\alpha)$ was expected. Therefore, for a chain growth probability of $\alpha = 0.9$, the slope was expected to be -0.04576 . The slope of the modelled products was found to be -0.4630 . Similar to [24], C_{30} was not included in the distribution plot as it is representative of the lumped components C_{21-30} . When comparing the simulated slope to the theoretical slope, it deviated by only 1.2 %. Based on the small deviation seen, the simulated products were deemed to be adequate.

3.4.7 RECYCLING

With the FTR section validated, it was possible to build the remainder of the GTL process. In order to yield the two streams that a multi-tubular fixed bed reactor would produce, a separator was added after the PFR, hence providing a gaseous product stream ([Stream 13](#)) and a liquid product stream ([Stream 14](#)). To remove some of the unwanted water in the vapour stream, it was cooled to 38 °C. This cooled stream and the liquid product stream were then fed into a three-phase separator. In an actual system, the light liquid products ([Stream 17](#)) and heavy liquid products ([Stream 18](#)) are usually forwarded to the upgrading section (not included in this study). The vapour products stream ([Stream 16](#)) is split into a recycle stream and purge stream (Using [Splitter 1](#) in a ratio of 0.8:0.2). In the model of [24], the recycled stream was compressed ([Compressor](#)) and fed back to the pre-reformer and FTR in a ratio of 0.232:0.768 ([Splitter 2](#)). Seeing as this study excluded the pre-reformer, the stream was purged ([Stream 22](#)). The connection of the *recycle block* concluded the simulation effort. The stream information is recapitulated in Table 3.14, and the complete HYSYS® process flow diagram is depicted in Figure 3.9.

Table 3.12: Summary of the weight fraction per carbon number (C_{1-14})

n	w_n	w_n/n	$\log(w_n/n)$	n	w_n	w_n/n	$\log(w_n/n)$
1	0.159781	0.159781	-0.79647	2	0.002998	0.001499	-2.82414
3	0.003957	0.001319	-2.87971	4	0.004701	0.001175	-2.92987
5	0.005276	0.001055	-2.97667	6	0.005633	0.000939	-3.02739
7	0.005884	0.000841	-3.07542	8	0.006075	0.000759	-3.11955
9	0.006110	0.000679	-3.16817	10	0.006148	0.000615	-3.21126
11	0.006061	0.000551	-3.25882	12	0.005850	0.000488	-3.31199
13	0.005719	0.000440	-3.35659	14	0.005495	0.000392	-3.40618

Table 3.13: Summary of the weight fraction per carbon number (C₁₅₋₂₀)

n	w_n	w_n/n	$\log(w_n/n)$	n	w_n	w_n/n	$\log(w_n/n)$
15	0.005413	0.000361	-3.44268	16	0.005268	0.000329	-3.48246
17	0.005062	0.000298	-3.52615	18	0.004793	0.000266	-3.57466
19	0.004462	0.000235	-3.62919	20	0.004382	0.000219	-3.65931

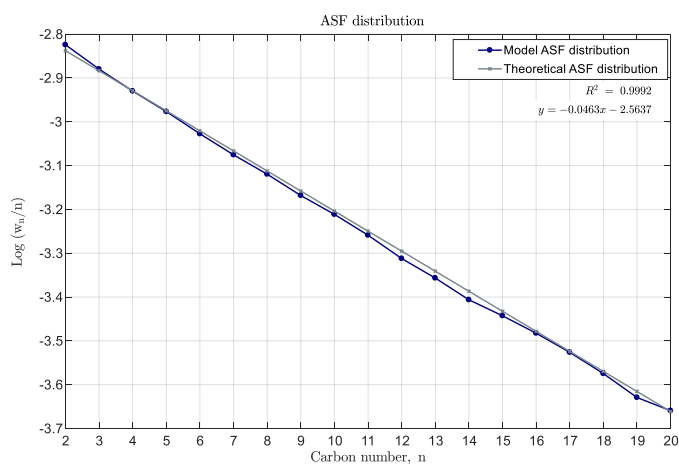
Figure 3.8: The ASF distribution of the FTR products C₂-C₂₀ (Stream 12)

Table 3.14: Stream information of the simulated GTL process

Stream no	Description	Temperature [°C]	Molar flow [kgmole/h]	Pressure [kPa]
Syngas production section				
1	Methane	675	8195.0	3000
2	Steam	675	5429.2	3000
3	Oxygen	200	4434.9	3000
4	Carbon dioxide	675	959.4	3000
5	Syngas	1029	30262.6	3000
7	Cooled syngas	38	30262.6	3000
8	Cleaned syngas	38	24452.7	3000
Fischer-Tropsch section				
10	Mixed stream 1	54	34310.3	3000
11	Reactor feed	210	34310.3	2000
12	Reactor products	213	22290.4	1940
13	Gaseous products	213	22238.3	1940
14	Liquid products	213	52.1	1940
15	Cooled reactor products	38	22238.3	1940
16	Vapour products	44	16063.3	1940
17	Light liquid products	44	189.8	1940
18	Heavy liquid products	44	6037.3	1940
Recycle section				
19	Purge 1	44	3212.7	1940
20	Recycle gas	44	12850.6	1940
21	Compressed gas	88	12850.6	3000
22	Purge 2	88	2981.3	3000
24	Recycle to FTR	88	9857.7	3000

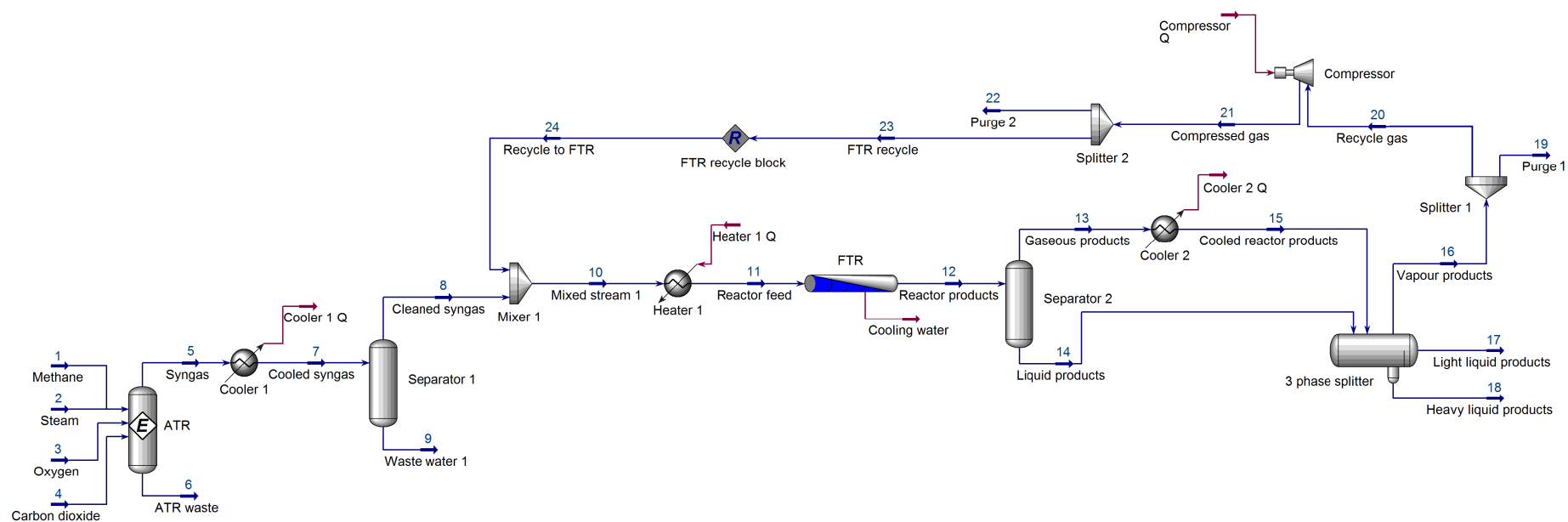


Figure 3.9: The Aspen HYSYS[®] process flow diagram of the developed GTL process

3.5 FAULT CONDITIONS

3.5.1 FAULT RATIONALE

In order to evaluate any proposed FDI technique, applicable fault conditions of the system are crucial. To determine the faults a GTL system might undergo, the major process units were identified. The most prominent units were found to be process control units, rotating equipment, heat transfer equipment, and reactors. The specific type of equipment categorised under each of these are diagrammatically summarised in Figure 3.10. When evaluating relevant literature, recurring causes of potential failures are well-documented. Although failures are defined as permanent interruptions, very few incidents are because of sudden and spontaneous failures. More often than not, the failures are a result of disregarded warnings and faults [1]. As such, the causes of failures could be used as a basis to define faults. Note that only the units that can easily be manipulated or imitated within the simulation environment, HYSYS®, are assessed. Usually, when specific faults or failures occur, a resultant effect(s) is observed within the unit and/or system. Table 3.15 summarises the causes and the effects that are seen within the considered GTL units.

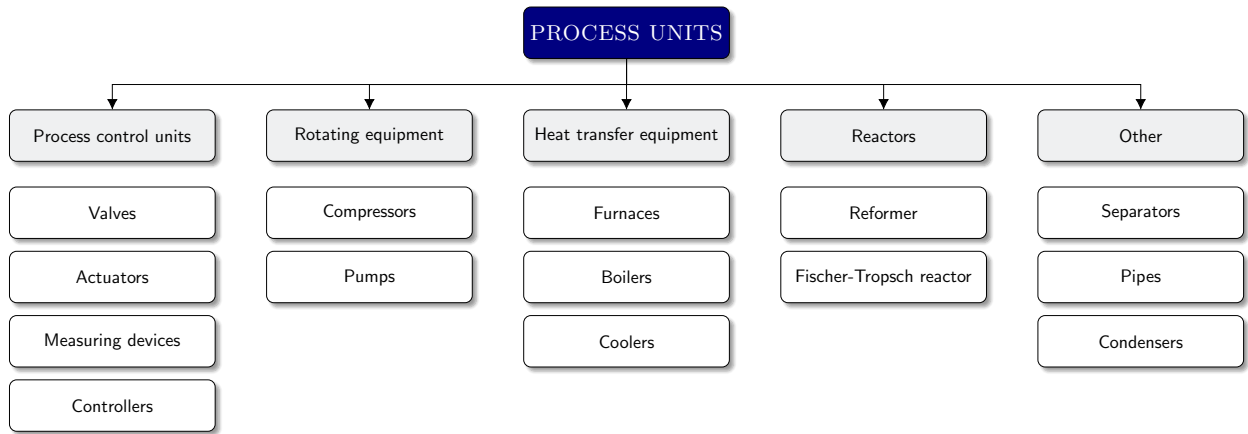


Figure 3.10: The prominent process units seen in a GTL process

3.5.2 FAULT SETS

All the considered faults were based on the above-detailed causes and effects. As mentioned, only faults that would be possible to emulate within HYSYS®, were considered. Additionally, not all possible causes were assimilated; rather, a representative set was deemed sufficient. For every considered fault, a fault ID was assigned in the form F_{pq_r} . In order to formulate the faults, the GTL process was firstly divided into its three central processing sections ($p = 1, 2$, and 3). For every unique type of fault per section ($q = 1, 2, \dots, 4$), seven different magnitudes were assigned ($r = 1, 2, \dots, 7$), the corresponding percentages shown in Table 3.16. Thus, a total of 84 faults were specified. The locations of the faults are schematically shown, using danger triangles, in Figure 3.11. The smallest fault magnitude was chosen as 3 %. This was to ensure that the effects of a small fault would be distinguishable from the solver deviations encountered (documented in Appendix D). The subsequent subsections will discuss in more detail what each section's fault set comprised of.

Table 3.15: Common recurring causes and effects of unit failures [1, 148]

Unit	Common cause	Effects
Valves	→ The incorrect type of valve installed in specific applications. → Insufficient lubrication/cleaning of valves and relevant parts. → Gasket defects/failures. → Spring defects/failures. → Incorrect/inadequate mounting of the valves. → Insufficient bleed-off tuning resulting in equipment being damaged.	→ Directly influences the flow rate of the stream. → Affects the pressure of the stream.
Actuators	→ Insufficient actuator capacity when there is no pipeline pressure assisting. → Actuator stem becomes corroded because of incorrect material selection.	→ Directly influences the flow rate of the stream. → Affects the pressure of the stream.
Rotating equipment	→ Seals', bearings' and bodies' degradation effects are increased if ○ incorrect installations occur ○ the equipment is maloperated ○ inadequately maintained ○ there are manufacturing flaws present	→ Affects the pressure of the stream. → Affects the flow rate of the stream.
Pipes	→ Fatigue → Inadequate flexibility → Corrosion/erosion effects → Insufficient support structures leaves pipes ○ free to vibrate ○ vulnerable to sagging when the weight of the transported material changes → Incorrect pipes were installed for the specific application → Poor maintenance practices are followed → Defects in the flanges/gaskets	→ Affects the flow rate of the stream. → Affects the pressure of the stream.
Heat exchanger (HE)	→ Excessive cooling can cause the metal to become brittle, leading to cracks which leaks (fluid leaks). → The build-up of debris on the heat transfer surface is the main cause of fouling. → Insufficient insulation can cause heat leakages to occur, decreasing the HE efficiency. → Water hammer (hydraulic shock) in the pipelines can damage various aspects of the heat exchanging workings	→ <i>Fouling</i> creates a pressure drop. → <i>Fluid leak</i> and <i>heat leakage</i> affects the efficiency of the HE which directly influences the streams' temperature.
Reactors	→ Fouling of the catalyst bed caused by contaminants.	→ <i>Fouling</i> of the catalyst bed can create a pressure drop.

Table 3.16: Percentage magnitudes of signifier r

r	1	2	3	4	5	6	7
Magnitude	3 %	8 %	9 %	10 %	11 %	12 %	25 %

3.5.2.1 FAULT SET F_{1q_r}

When specifying the faults within the syngas production section (F_{1q_r}), the following effects were of particular interest: The same type of fault of the same magnitude and location but in opposite directions. These faults are notated as F_{11_r} and F_{12_r} respectively, and are representative of deviations in the feed molar flow rate (caused by either faulty valves or actuators). Two different types of faults of the same magnitude and location. Here referring to F_{12_r} and F_{13_r} . Fault F_{13_r} is the result of the feed stream's pressure being too low. The same type of fault of the same magnitude but slightly different locations, F_{13_r} and F_{14_r} . Fault F_{14_r} would represent fouling of the reactor bed. Table 3.17 summarises this section's considered faults and corresponding particulars.

Table 3.17: The location and details of simulated faults F_{1q_r}

F_{1q_r}		Syngas production section		
Fault ID	Location	Description	Details	
F_{11_r}	F_{111}	Methane stream	Molar flow +3 %	+245.9 kgmole/h
	F_{112}	Methane stream	Molar flow +8 %	+655.6 kgmole/h
	F_{113}	Methane stream	Molar flow +9 %	+737.6 kgmole/h
	F_{114}	Methane stream	Molar flow +10 %	+819.5 kgmole/h
	F_{115}	Methane stream	Molar flow +11 %	+901.5 kgmole/h
	F_{116}	Methane stream	Molar flow +12 %	+983.4 kgmole/h
	F_{117}	Methane stream	Molar flow +25 %	+2048.75 kgmole/h
F_{12_r}	F_{121}	Methane stream	Molar flow -3 %	-245.9 kgmole/h
	F_{122}	Methane stream	Molar flow -8 %	-655.6 kgmole/h
	F_{123}	Methane stream	Molar flow -9 %	-737.6 kgmole/h
	F_{124}	Methane stream	Molar flow -10 %	-819.5 kgmole/h
	F_{125}	Methane stream	Molar flow -11 %	-901.5 kgmole/h
	F_{126}	Methane stream	Molar flow -12 %	-983.4 kgmole/h
	F_{127}	Methane stream	Molar flow -18 %	-1475.1 kgmole/h
F_{13_r}	F_{131}	Methane stream	Pressure -3 %	-90 kPa
	F_{132}	Methane stream	Pressure -8 %	-240 kPa
	F_{133}	Methane stream	Pressure -9 %	-270 kPa
	F_{134}	Methane stream	Pressure -10 %	-300 kPa
	F_{135}	Methane stream	Pressure -11 %	-330 kPa
	F_{136}	Methane stream	Pressure -12 %	-360 kPa
	F_{137}	Methane stream	Pressure -25 %	-750 kPa
F_{14_r}	F_{141}	ATR	Pressure -3 %	-90 kPa
	F_{142}	ATR	Pressure -8 %	-240 kPa
	F_{143}	ATR	Pressure -9 %	-270 kPa
	F_{144}	ATR	Pressure -10 %	-300 kPa
	F_{145}	ATR	Pressure -11 %	-330 kPa
	F_{146}	ATR	Pressure -12 %	-360 kPa
	F_{147}	ATR	Pressure -25 %	-750 kPa

3.5.2.2 FAULT SET F_{2q_r}

Similarly, for the Fischer-Tropsch section (F_{2q_r}), the subsequent faults were evaluated. As the Fischer-Tropsch process is sensitive to deviations in temperature; fault F_{21_r} is the result of insufficient heating of [Heater 1](#), delivering reactor feed at a lower than expected temperature. F_{24_r} is representative of a problem regarding the water cooling of the reactor, causing the reaction temperature to increase. F_{22_r} and F_{23_r} are based on the notion that there could be damaged pipes, resulting in leakages and pressure drops. The details of these faults are outlined in Table 3.18

Table 3.18: The location and details of simulated faults F_{2q_r}

F_{2q_r}		Fischer-Tropsch section			
Fault ID	Location	Description	Details		
F_{21_r}	F_{21_1}	Reactor feed stream	Temperature -3%	$-6.3\text{ }^{\circ}\text{C}$	$203.7\text{ }^{\circ}\text{C}$
	F_{21_2}	Reactor feed stream	Temperature -8%	$-16.8\text{ }^{\circ}\text{C}$	$193.2\text{ }^{\circ}\text{C}$
	F_{21_3}	Reactor feed stream	Temperature -9%	$-18.9\text{ }^{\circ}\text{C}$	$191.1\text{ }^{\circ}\text{C}$
	F_{21_4}	Reactor feed stream	Temperature -10%	$-21.0\text{ }^{\circ}\text{C}$	$189.0\text{ }^{\circ}\text{C}$
	F_{21_5}	Reactor feed stream	Temperature -11%	$-23.1\text{ }^{\circ}\text{C}$	$186.9\text{ }^{\circ}\text{C}$
	F_{21_6}	Reactor feed stream	Temperature -12%	$-25.2\text{ }^{\circ}\text{C}$	$184.8\text{ }^{\circ}\text{C}$
	F_{21_7}	Reactor feed stream	Temperature -25%	$-52.5\text{ }^{\circ}\text{C}$	$157.5\text{ }^{\circ}\text{C}$
F_{22_r}	F_{22_1}	Reactor feed stream	Leakage -3%	Splitter 0.97:0.03	N/A
	F_{22_2}	Reactor feed stream	Leakage -8%	Splitter 0.92:0.08	N/A
	F_{22_3}	Reactor feed stream	Leakage -9%	Splitter 0.91:0.09	N/A
	F_{22_4}	Reactor feed stream	Leakage -10%	Splitter 0.90:0.10	N/A
	F_{22_5}	Reactor feed stream	Leakage -11%	Splitter 0.89:0.11	N/A
	F_{22_6}	Reactor feed stream	Leakage -12%	Splitter 0.88:0.12	N/A
	F_{22_7}	Reactor feed stream	Leakage -25%	Splitter 0.75:0.25	N/A
F_{23_r}	F_{23_1}	FTR	Pressure -3%	-60 kPa	1940 kPa
	F_{23_2}	FTR	Pressure -8%	-160 kPa	1840 kPa
	F_{23_3}	FTR	Pressure -9%	-180 kPa	1820 kPa
	F_{23_4}	FTR	Pressure -10%	-200 kPa	1800 kPa
	F_{23_5}	FTR	Pressure -11%	-220 kPa	1780 kPa
	F_{23_6}	FTR	Pressure -12%	-240 kPa	1760 kPa
	F_{23_7}	FTR	Pressure -25%	-500 kPa	1500 kPa
F_{24_r}	F_{24_1}	FTR	Temperature -3%	$-3.39\text{E}+07\text{ kJ/h}$	1096100000 kJ/h
	F_{24_2}	FTR	Temperature -8%	$-9.04\text{E}+07\text{ kJ/h}$	1039600000 kJ/h
	F_{24_3}	FTR	Temperature -9%	$-1.02\text{E}+08\text{ kJ/h}$	1028000000 kJ/h
	F_{24_4}	FTR	Temperature -10%	$-1.13\text{E}+08\text{ kJ/h}$	1017000000 kJ/h
	F_{24_5}	FTR	Temperature -11%	$-1.24\text{E}+08\text{ kJ/h}$	1006000000 kJ/h
	F_{24_6}	FTR	Temperature -12%	$-1.36\text{E}+08\text{ kJ/h}$	994000000 kJ/h
	F_{24_7}	FTR	Temperature -25%	$-2.83\text{E}+08\text{ kJ/h}$	847000000 kJ/h

3.5.2.3 FAULT SET F_{3q_r}

Finally, the recycle section (F_{3q_r}) was subjected to the following faults. The recycle compressor could degrade over time, resulting in lower compression, F_{31_r} being achieved. A blockage in the gas splitter F_{32_r} causes less gas being recycled and a subsequent increase in purge gas volume. Likewise, F_{33_r} , a blockage would cause a higher ratio of gas to be recycled to the FTR. F_{34_r} would simulate the effect of a pipe leak in

the recycle stream itself. The primary purpose of investigating faults within the recycle stream is to evaluate whether the fault location can be pinpointed or whether it will inevitably propagate throughout the entire process. Table 3.19 depicts the particulars of the recycle section faults.

Table 3.19: The location and details of simulated faults F_{3q_r} .

F_{3q_r}		Recycle section			
Fault ID	Location	Description	Details		
F_{31_r}	F_{31_1}	Compressor	Pressure -3%	-90 kPa	2910 kPa
	F_{31_2}	Compressor	Pressure -8%	-240 kPa	2760 kPa
	F_{31_3}	Compressor	Pressure -9%	-270 kPa	2730 kPa
	F_{31_4}	Compressor	Pressure -10%	-300 kPa	2700 kPa
	F_{31_5}	Compressor	Pressure -11%	-330 kPa	2670 kPa
	F_{31_6}	Compressor	Pressure -12%	-360 kPa	2640 kPa
	F_{31_7}	Compressor	Pressure -25%	-750 kPa	2250 kPa
F_{32_r}	F_{32_1}	Splitter 1	Lower split ratio -3%	$0.776:0.224$	N/A
	F_{32_2}	Splitter 1	Lower split ratio -8%	$0.736:0.264$	N/A
	F_{32_3}	Splitter 1	Lower split ratio -9%	$0.728:0.272$	N/A
	F_{32_4}	Splitter 1	Lower split ratio -10%	$0.720:0.280$	N/A
	F_{32_5}	Splitter 1	Lower split ratio -11%	$0.712:0.288$	N/A
	F_{32_6}	Splitter 1	Lower split ratio -12%	$0.704:0.296$	N/A
	F_{32_7}	Splitter 1	Lower split ratio -25%	$0.600:0.400$	N/A
F_{33_r}	F_{33_1}	Splitter 2	Higher split ratio $+3\%$	$0.7910:0.2090$	N/A
	F_{33_2}	Splitter 2	Higher split ratio $+8\%$	$0.8294:0.1706$	N/A
	F_{33_3}	Splitter 2	Higher split ratio $+9\%$	$0.8371:0.1629$	N/A
	F_{33_4}	Splitter 2	Higher split ratio $+10\%$	$0.8448:0.1552$	N/A
	F_{33_5}	Splitter 2	Higher split ratio $+11\%$	$0.8525:0.1475$	N/A
	F_{33_6}	Splitter 2	Higher split ratio $+12\%$	$0.8602:0.1398$	N/A
	F_{33_7}	Splitter 2	Higher split ratio $+25\%$	$0.9600:0.0400$	N/A
F_{34_r}	F_{34_1}	Recycle to FTR	Leakage -3%	Splitter $0.97:0.03$	N/A
	F_{34_2}	Recycle to FTR	Leakage -8%	Splitter $0.92:0.08$	N/A
	F_{34_3}	Recycle to FTR	Leakage -9%	Splitter $0.91:0.09$	N/A
	F_{34_4}	Recycle to FTR	Leakage -10%	Splitter $0.90:0.10$	N/A
	F_{34_5}	Recycle to FTR	Leakage -11%	Splitter $0.89:0.11$	N/A
	F_{34_6}	Recycle to FTR	Leakage -12%	Splitter $0.88:0.12$	N/A
	F_{34_7}	Recycle to FTR	Leakage -25%	Splitter $0.75:0.25$	N/A

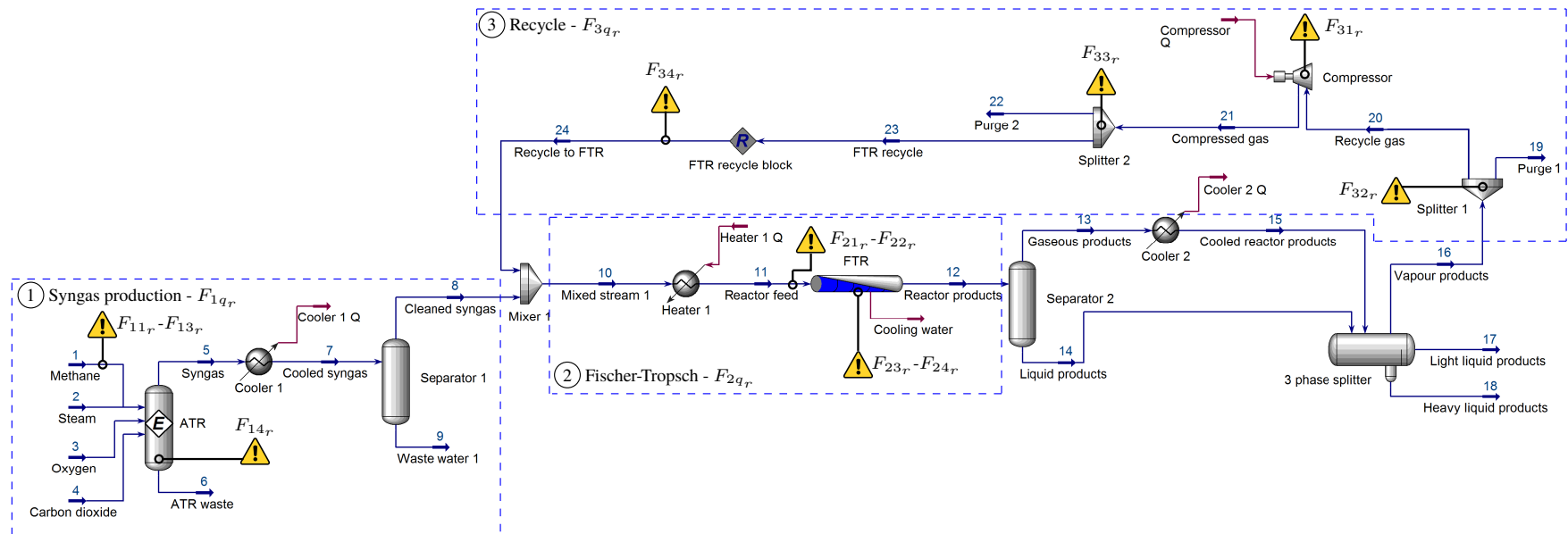


Figure 3.11: The GTL process showing the considered fault locations

3.6 CONCLUSION

Synthetic fuel has long been an essential part of many economies, and as naturally-occurring petroleum reserves dwindle, it will remain as such. A popular route to produce synthetic fuels is the GTL process which utilises natural gas as feedstock. Seeing as the GTL process would also be ideally suited as a basis for an FDI study, a simulation model of representative complexity and scale was developed. The model was simulated in Aspen HYSYS® and validated using the product distribution seen within literature. Special care was taken to document the modelling specifics of the GTL simulation so that the model might be used as an alternative larger-scale FDI benchmark system. Using common failures seen in various process units, plausible fault conditions were defined. With a viable simulation model, the characterisation of the process in terms of energy properties can now be undertaken. Chapter 4 will detail some essential theoretical fundamentals related to exergy and the quantification thereof as well as the manner in which the GTL is characterised.

CHAPTER 4

ENERGY CHARACTERISATION

4.1 INTRODUCTION

Chapter 3 showed the workings and development of the GTL model that will be used throughout this study. As the energy-based FDI approaches make use of energy properties, the GTL process and all of its constituents need to be described in terms of physical and chemical exergy as well as energy flow between connected process units. This chapter stipulates how the energy properties were calculated. The all-important reference environment (RE) is firstly defined. Next, the general calculation of physical and chemical exergy is shown. In order to automatically compute the considered exergies, *user variables* were developed within HYSYS®. A *user variables* is Visual Basic for Applications (VBA) code that a user can create to access and manipulate various components of the HYSYS® model. The particulars surrounding the development of the *user variables* are comprehensively shown and discussed. Additionally, the validation of the exergy values obtained is included. The energy flows that will be used are specified and the chapter is concluded by emphasising some important fault datasets.

4.2 BACKGROUND TO EXERGY

According to [25], exergy is defined as being a quantitative measure of an energy quantity's usefulness to perform work. In other words, the maximum theoretical useful work obtained if a system is brought to thermodynamic equilibrium with the environment by means of processes which the system only interacts with the environment. Unlike energy which is based only on the first law of thermodynamics, exergy also takes into account the second law of thermodynamics. The second law states that entropy cannot decrease in any real process, therefore the ability to deliver valuable work is eventually lost. Simply put, exergy is not conserved and some exergy losses would occur which could be quantified by using the process' exergy balance [149]. The most prominent advantage of using exergy is, therefore, that it enables a manner of quantifying the quality of an energy stream or (more importantly) the efficiency of various elements. Consequently, any deviation of the known efficiencies could be indicative of an anomaly within the system. Fundamentally, the use of exergy as the monitored parameter leverage the structural information contained in the process itself [150]. It has also been shown that exergetic efficiencies could also be used to diagnose the component level under

performance in a biomass gassifier [151], and similar work was done by [152, 153] pertaining to turbines. This seems to suggest that exergy is well suited to detect component level inefficiencies (degradation or faults) when the system level performance degrades. Indeed, [154] showed that exergy can be used to determine the efficiency of fired heaters, of which the ATR is a typical example. For many PCI processes, the cyclic nature of the process presents a particularly challenging scenario as feedstocks and products are cycled, recycled, and discarded.

4.3 EXERGY CALCULATIONS

4.3.1 REFERENCE ENVIRONMENT

Exergy is always evaluated relative to a reference environment (RE). This means that the RE's intensive properties will determine the exergy. A reference environment is defined as an infinite system which is in stable equilibrium, where all parts thereof are at rest to one another. Moreover, no chemical reactions between its environmental components can occur. It can only witness internally reversible processes where its own intensive properties remain unchanged and the chemical potentials of constituents stay constant. The natural environment, however, does not exhibit the same characteristics as a theoretical environment; as it is not in equilibrium and the intensive properties demonstrate changes. Therefore, models for reference environments usually aim to incorporate some of the natural environment's behaviour alongside the theoretical aspects [25]. A few different RE-models were developed over the years, with the most prominent ones being the work of Van Gool [155], Szargut [156], Wepfer and Gaggioli [157], and Ahrendts [158]. In the study done by Muñoz et al. [159] the impact of the reference environment (RE) on exergy analyses were investigated. It was shown that when working with different RE-models, varying chemical exergy results are obtained when considering absolute exergies. Conversely, when observing exergy destruction or efficiencies, the different models yielded comparable results. According to [158], however, the model proposed by Szargut [156] is utilised in most engineering applications as it delivers appropriate and consistent results. As such, this study will also make use of Szargut's model.

4.3.2 TOTAL EXERGY

The total exergy of a system, which has no magnetic, nuclear, electric, or surface tension characteristics, is usually expressed as:

$$b_{tot} = b_{kin} + b_{pot} + b_{ph} + b_{ch}. \quad (4.1)$$

Here b_{kin} refers to the kinetic exergy, b_{pot} to the potential exergy, b_{ph} to the physical exergy and b_{ch} to the chemical exergy. As the actual GTL plant is static, the kinetic and potential exergy can be disregarded, and (4.1) is simplified to

$$b_{tot} = b_{ph} + b_{ch}. \quad (4.2)$$

The subsequent sections will document the details and calculations of physical exergy and chemical exergy, respectively.

4.3.3 PHYSICAL EXERGY

4.3.3.1 THEORY

The physical exergy is the work that can be obtained by taking the system from its initial state (T and P), at a constant composition, to the considered reference environment's (RE) temperature (T_0) and pressure (P_0). These are generally defined as $T_0 = 25\text{ }^\circ\text{C}$ and $P_0 = 101.325\text{ kPa}$ and are used as such within this study. Therefore, to calculate the physical exergy for 1 mole of constituent,

$$b_{ph} = (h - h_0) - T_0(s - s_0), \quad (4.3)$$

is implemented; with h referring to the enthalpy and s the entropy of the initial state. Whereas h_0 and s_0 refers to the reference environment's enthalpy and entropy at temperature T_0 . To obtain the total physical exergy (B_{ph}), the calculated b_{ph} is multiplied by the total molar flow (n) of the considered stream. Mathematically, this is expressed as:

$$B_{ph} = b_{ph}n. \quad (4.4)$$

4.3.3.2 USER VARIABLE

In order to implement (4.4) in HYSYS[®], a *user variable* was developed based on Algorithm 1. The algorithm starts by obtaining the stream's current enthalpy and entropy. Next the stream's temperature and pressure are set to that of the reference environment. It then forces a recalculation of the stream's enthalpy and entropy. The physical exergy per mole is then computed by implementing (4.4). The stream's total physical exergy (B_{ph}) is lastly obtained by multiplying the per mole exergy by the stream's molar flow. The verbatim VBA code is given in Appendix C. It should be noted that the *user variable* was applied to every stream within the GTL simulation.

Algorithm 1 Computing a stream's physical exergy (B_{ph})

Require: Reference environment temperature T_0 and pressure P_0 in the simulation

```
1: Stream ← ActiveObject.DuplicateFluid
2: if Stream.MolarFlow.IsKnown
3: and Stream.MolarFractions.IsKnown
4: and Stream.VapourFraction.IsKnown
5: and Stream.Pressure.IsKnown then                                ▷ conditions should be known
6:    $h \leftarrow \text{Stream.MolarEnthalpy.GetValue}(\text{"kJ/kgmole"})$         ▷ obtain current stream enthalpy
7:    $s \leftarrow \text{Stream.MolarEntropy.GetValue}(\text{"kJ/kgmole-C"})$       ▷ obtain current stream entropy
8:    $\text{Stream.Temperature} \leftarrow T_0$                                 ▷ set stream temperature to reference  $T_0$ 
9:    $\text{Stream.Pressure} \leftarrow P_0$                                     ▷ set stream pressure to reference  $P_0$ 
10:   $\text{Stream.TPFlash}()$                                                 ▷ PV flash is executed
11:   $h_0 \leftarrow \text{Stream.MolarEnthalpy.GetValue}(\text{"kJ/kgmole"})$       ▷ obtain stream enthalpy after flash
12:   $s_0 \leftarrow \text{Stream.MolarEntropy.GetValue}(\text{"kJ/kgmole-C"})$     ▷ obtain stream entropy after flash
13:   $B_{ph} \leftarrow (h - h_0) - (T_0 + 273.15)(s - s_0)$               ▷ calculate physical exergy per mole
14:   $F \leftarrow \text{Stream.MolarFlow.GetValue}(\text{"kgmole/h"})$             ▷ obtain molar flow rate of stream
15:   $B_{ph} \leftarrow B_{ph}F$                                           ▷ multiply per mole exergy with molar flow rate to obtain total physical exergy
16: end if
```

4.3.3.3 VALIDATION

To validate the *user variable* calculations, the values were compared to hand calculations. The hand calculation, depicted in (4.5), uses parameters that are automatically calculated within HYSYS®. The values obtained from HYSYS® and the computational results are shown in Table 4.1. Based on the insignificant differences seen, the *user variable* calculations were regarded acceptable.

$$B_{ph} = \text{Mass flow} \times \text{Mass exergy} \quad (4.5)$$

Table 4.1: Hand calculated and user variable exergy values for the methane stream (Stream 1) compared

Stream no.	Mass flow [kg/h]	Mass exergy [kJ/kg]	Hand calculated [kJ/h]	User variable [kJ/h]	Difference [%]
1	131447.8	1612.7	211985867	211991000	0.002

4.3.4 CHEMICAL EXERGY

4.3.4.1 THEORY

As mentioned, the reference environment comprises of certain reference elements and intensive properties. The RE chosen as basis, is the one proposed by Szargut [156, 160]. Szargut suggests that the chemical exergy obtained in the standard state at normal temperature and pressure conditions should be considered as a *standard chemical exergy* (b_{ch}^0); the environment consisting of a set of reference elements with standard concentrations based on conventional means. By having the standard chemical exergy values of elements, the chemical exergy of any chemical compound can be determined by utilising

$$b_{ch_n}^0 = \bar{g}^0 + \sum_e n_e b_{ch_e}^0. \quad (4.6)$$

Here $\bar{g}^0$ is the Gibbs free energies of formation, n_e the amount of substance n and $b_{ch_e}^0$ the standard chemical exergy of substance n .

4.3.4.2 USER VARIABLE

To calculate the chemical exergy within HYSYS®, the following two points are of importance. Firstly, as detailed in [16], HYSYS® takes exergy of mixing into account. Thus, the expression for chemical exergy of mixtures:

$$b_{ch} = \sum x_{(i)} b_{ch_{(i)}}^0 + RT_0 \sum x_{(i)} \ln x_{(i)}, \quad (4.7)$$

can be simplified to

$$b_{ch} = \sum x_{(i)} b_{ch_{(i)}}^0. \quad (4.8)$$

When employing (4.8) then, one requires only the mole fraction $x_{(i)}$ and standard chemical exergy $b_{ch_{(i)}}^0$ of substance i . For a GTL process, there will understandably be multi-phase streams. Some substances, such

as water, have different standard chemical exergy values when in different phases. To take this into account, (4.8) is extended; the total chemical exergy was taken as the sum of the vapour, the liquid, and the aqueous phase exergy. Mathematically this is conveyed as:

$$b_{ch} = \sum x_{(i)v} b_{ch(i)v}^0 + x_{(i)\ell} b_{ch(i)\ell}^0 + x_{(i)a} b_{ch(i)a}^0. \quad (4.9)$$

By multiplying the phase intrinsic chemical exergy with the relevant stream's total molar flow (n) the total chemical exergy (B_{ch}) is obtained. Equation (4.9) is, therefore, simply modified to become:

$$B_{ch} = \sum x_{(i)v} b_{ch(i)v}^0 n + x_{(i)\ell} b_{ch(i)\ell}^0 n + x_{(i)a} b_{ch(i)a}^0 n. \quad (4.10)$$

Secondly, as contended by [16], from a computational point of view it would be more efficient to make use of a look-up table to obtain standard chemical exergies already well-defined than having to recalculate them as done in the work of [161]. Therefore, the reference substances and their corresponding standard chemical exergy (taken from [160]) were placed in a *user property* in order to be accessible by the *simulation basis*. Some of the hydrocarbons' standard chemical exergy were not readily available and had to be computed before being implemented within the *user property*. Appendix B documents the steps taken to calculate these. The chemical exergy *user variable* was developed based on Algorithm 2. A *user variable* for every phase was created and summed to determine the total chemical exergy of the considered stream. Looking at the vapour phase *user variable*; it starts off by obtaining the component's vapour molar flow (m_v) and total molar flow (m_T). Dividing these values, the component's vapour ratio is found $\text{ratio}_v = \frac{m_v}{m_T}$. In order to determine the component's vapour phase mole fraction (mF_v), the component's total mole fraction is multiplied by the vapour ratio ($mF_T \cdot \text{ratio}_v$). Next, the component's corresponding standard chemical exergy is extracted from the *user property*. This is then multiplied with the vapour phase mole fraction (mF_v) and total molar flow (m_T) to obtain the component's vapour chemical exergy (B_{ch_v}). The stream constituents' chemical exergy is summed to finally produce the total chemical exergy (B_{ch}), for the vapour phase. The liquid and aqueous phases were conducted in a similar manner.

4.3.4.3 VALIDATION

As with the physical exergy, hand calculations were used to validate the user variable's values. Table 4.2 shows the simulated mole fractions along with the standard chemical exergy (vapour phase) of the syngas stream ([Stream 5](#)) components. Using (4.8) and summing the obtained chemical exergy values, the comparison in Table 4.3 depicts adequate results.

Table 4.2: Hand calculated chemical exergy for components in syngas stream ([Stream 5](#))

Stream composition		$b_{ch(i)}^0$	Molar flow rate	Calculated exergy
Substance	Mole fraction	[kJ/kgmole]	[kgmole/h]	[kJ/h]
H ₂ O(g)	0.1940	9500	30262.6	55763999
CO	0.2396	274710	30262.6	1991899948
CO ₂	0.0512	19480	30262.6	30183186
CH ₄	0.0117	831200	30262.6	295177811
H ₂	0.5035	236090	30262.6	3597355057
O ₂	0.0000	3970	30262.6	0

Table 4.3: Hand calculated and user variable exergy values for the syngas stream ([Stream 5](#)) compared

Stream no.	Hand calculated [kJ/h]	User variable [kJ/h]	Difference [%]
5	5970380000	5970380000	0.0000

Algorithm 2 Computing a stream's chemical exergy (B_{ch})

Require: Standard chemical exergy (B_{ch}^0) stored in a *user property* for each component (phase specific)

```
1: Stream  $\leftarrow$  ActiveObject.DuplicateFluid
2: if Stream.Pressure.IsKnown
3: and Stream.VapourFraction.IsKnown
4: and Stream.MolarFlow.IsKnown
5: and Stream.MolarFractions.IsKnown then ▷ conditions should be known
6:   Components  $\leftarrow$  Stream.Components ▷ components present within stream
7:    $B_{ch} = 0$ 
8:   for each Component do
9:     if Stream.MolarFlows.Values > 0 then
10:      if phase  $\leftarrow$  vapour then
11:         $m_v \leftarrow$  Stream.VapourPhase.MolarFlows.Values ▷ component's vapour molar flow
12:         $m_T \leftarrow$  Stream.MolarFlows.Values ▷ component's total molar flow
13:         $ratio_v = m_v / m_T$ 
14:         $mF_T \leftarrow$  Stream.MolarFractionsValue ▷ component's total mole fraction
15:         $mF_v = ratio_v * mF_T$  ▷ component's vapour phase mole fraction
16:         $B_{ch_v}^0 \leftarrow$  Component.GetUserProperty() ▷ component's vapour std chemical exergy
17:         $B_{ch_v} = mF_v * B_{ch_v}^0 * m_T$  ▷ compute component vapour chemical exergy
18:         $B_{ch} = B_{ch} + B_{ch_v}$  ▷ compute total vapour chemical exergy
19:      else if phase  $\leftarrow$  liquid then
20:         $m_l \leftarrow$  Stream.LightLiquidPhase.MolarFlows.Values ▷ liquid molar flow
21:         $m_T \leftarrow$  Stream.MolarFlows.Values ▷ component's total molar flow
22:         $ratio_l = m_l / m_T$ 
23:         $mF_T \leftarrow$  Stream.MolarFractionsValue ▷ component's total mole fraction
24:         $mF_l = ratio_l * mF_T$  ▷ component's liquid phase mole fraction
25:         $B_{ch_l}^0 \leftarrow$  Component.GetUserProperty() ▷ component's liquid std chemical exergy
26:         $B_{ch_l} = mF_l * B_{ch_l}^0 * m_T$  ▷ compute component liquid chemical exergy
27:         $B_{ch} = B_{ch} + B_{ch_l}$  ▷ compute total liquid chemical exergy
28:      else
29:         $m_a \leftarrow$  Stream.HeavyLiquidPhase.MolarFlows.Values ▷ aqueous molar flow
30:         $m_T \leftarrow$  Stream.MolarFlows.Values ▷ component's total molar flow
31:         $ratio_a = m_a / m_T$ 
32:         $mF_T \leftarrow$  Stream.MolarFractionsValue ▷ component's total mole fraction
33:         $mF_a = ratio_a * mF_T$  ▷ component's aqueous phase mole fraction
34:         $B_{ch_a}^0 \leftarrow$  Component.GetUserProperty() ▷ component's aqueous std chemical exergy
35:         $B_{ch_a} = mF_a * B_{ch_a}^0 * m_T$  ▷ compute component aqueous chemical exergy
36:         $B_{ch} = B_{ch} + B_{ch_a}$  ▷ compute total aqueous chemical exergy
37:      end if
38:    end if
39:  end for
40: end if
```

4.4 ENERGY CHARACTERISATION

The energy characterisation of a process stream consists of the physical exergy (B_{ph}), the chemical (B_{ch}) exergy, and energy flow ($\dot{E}$). As mentioned, the *user variables* are used to compute the desired exergy. To obtain the energy flow of the streams, the corresponding mass flow and mass enthalpy values, that are readily available within the HYSYS® workspace, are multiplied. As HYSYS® encounters small solver variations between runs, the normal base condition simulation model was run ten times to obtain the average energy quantities. The average normal condition is hereon referred to as the NOC. The ten normal conditions are labelled Normal₁ - Normal₁₀, and their data are tabulated in Appendix D. Table 4.4 shows the energy characterisation of the NOC per stream.

Table 4.4: Physical exergy, chemical exergy, and energy flow of each stream of the NOC

Stream no	Description	Exergy [kJ/h]		Energy flow $\dot{E}$ [kJ/h]
		B_{ph}	B_{ch}	
1	Methane	211991000	6811680000	-329629228
2	Steam	153042000	51577400	-1185124908
3	Oxygen	42119200	17608900	22860070
4	Carbon dioxide	22166700	18697700	-348963640
5	Syngas	853042500	5970380000	-1840857024
7	Cooled syngas	207502500	5919913100	-3102596564
8	Cleaned syngas	205658500	5915085000	-1445481368
10	Mixed stream 1	288901400	11343900000	-3032884567
11	Reactor feed	300436700	11343900000	-2852584935
12	Reactor products	222734000	10664647698	-3978187110
13	Gaseous products	220438400	9800699000	-3950984792
14	Liquid products	2178758	864081324	-27203643.6
15	Cooled reactor products	117531600	9748519204	-4378237273
16	Vapour products	116458300	8837496000	-2607799760
17	Light liquid products	246986	1769064000	-78084796.8
18	Heavy liquid products	533380	5608290	-1719558002
19	Purge 1	23291660	1767499000	-521559765
20	Recycle gas	93166650	7069996000	-2086238813
21	Compressed gas	108849400	7069996000	-2066399976
22	Purge 2	25253070	1640237000	-479404859
24	Recycle to FTR	83599810	5428814000	-1587406162

As 84 faults were too many to evaluate per approach, representative subsets were selected. The three prominent datasets, shown highlighted in Table 4.5, that are used throughout this study are:

- F_{pq_1} which includes the twelve faults that deviated with a magnitude of 3 %.
- F_{pq_4} which includes the twelve faults that deviated with a magnitude of 10 % as well as the NOC.
- F_{pq_R} which includes a random selection of various magnitude deviations - **excluding** 3 % and 10 % magnitudes - of each one of the twelve faults and the normal condition, Normal₁.

Table 4.5: Important fault datasets used in this study

Fault ID	Description	<i>r</i>								Location
		1	2	3	4	5	6	7		
<i>F</i> _{1<i>q_r</i>}		ATR section								
<i>F</i> _{11<i>r</i>}	Molar flow	+	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Methane stream
<i>F</i> _{12<i>r</i>}	Molar flow	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Methane stream
<i>F</i> _{13<i>r</i>}	Pressure	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Methane stream
<i>F</i> _{14<i>r</i>}	Pressure	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	ATR
<i>F</i> _{2<i>q_r</i>}		FTR section								
<i>F</i> _{21<i>r</i>}	Temperature	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Reactor feed stream
<i>F</i> _{22<i>r</i>}	Leakage	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Reactor feed stream
<i>F</i> _{23<i>r</i>}	Pressure	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	FTR
<i>F</i> _{24<i>r</i>}	Temperature	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	FTR
<i>F</i> _{3<i>q_r</i>}		Recycle section								
<i>F</i> _{31<i>r</i>}	Pressure	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Compressor
<i>F</i> _{32<i>r</i>}	Lower split ratio	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Splitter 1
<i>F</i> _{33<i>r</i>}	Higher split ratio	+	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Splitter 2
<i>F</i> _{34<i>r</i>}	Leakage	−	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Recycle to FTR

4.5 CONCLUSION

The reader was taken through some elementary theory related to exergy and the calculation thereof. The computation of the exergy of the streams within the GTL process was accomplished by exploiting *user variables*. The exact setup, code, and validation of these *user variables* were also shown. Subsequently, the exergy and energy flow data of every GTL stream and every individual fault condition were recorded. As 84 faults were too many to assess, representative subsets were identified. The energy data of these subsets can now be employed, as needed, within the energy-based FDI approaches. The first FDI approach to be examined, documented in the next chapter, is the exergy-based fixed-threshold approach.

CHAPTER 5

EXERGY-BASED FAULT DETECTION: A THRESHOLD APPROACH

5.1 INTRODUCTION

In the work done by Marais [16], the physical and chemical exergy of an Autothermal Reformer (ATR) were used in conjunction with a threshold function as means to the successful FDI of various faults. Advantages highlighted by the work were the use of energy, which is a unifying parameter across different domains as well as that it facilitates abstraction of data. Moreover, Marais emphasised that the applicability of the approach needed to be evaluated when applied to (1) a larger-scale plant and that (2) contains a recycle stream. Using some of Marais' work as a base, the proposed approach was applied to the GTL process which is compliant with both identified lacking properties. The exergy characteristics from Chapter 4 were utilised, along with an appropriate threshold function. The results are then interpreted in terms of detectability, isolability, and isolation performance. Lastly, the suitability of the approach is deliberated and discussed based on the performance metrics.

5.2 METHODOLOGY

In order to apply a threshold function to the considered fault data, a concise and repeatable methodology was required. The first subsection is used to give a quick overview of the threshold approach. The next subsections go into more detail on what the threshold and its constituent aspects are, how it was applied to the GTL data and the manner in which it was interpreted. Lastly, the evaluation of the approach's performance is discussed.

5.2.1 QUICK OVERVIEW

The threshold approach henceforth referred to as Approach I.A¹, is a qualitative approach which utilises the exergy characterisation of the GTL process from Chapter 4, along with a simplistic threshold function

¹Regarding the approach numbering, the A signify a qualitative approach, whereas B refers to a quantitative approach.

in order to obtain qualitative matrices. The qualitative matrices are then interpreted, the detail of which is given in Section 5.2.3, to determine fault detectability, isolability, and isolation. The reader is encouraged to refer to Figure 5.1 which summarises the approach graphically.

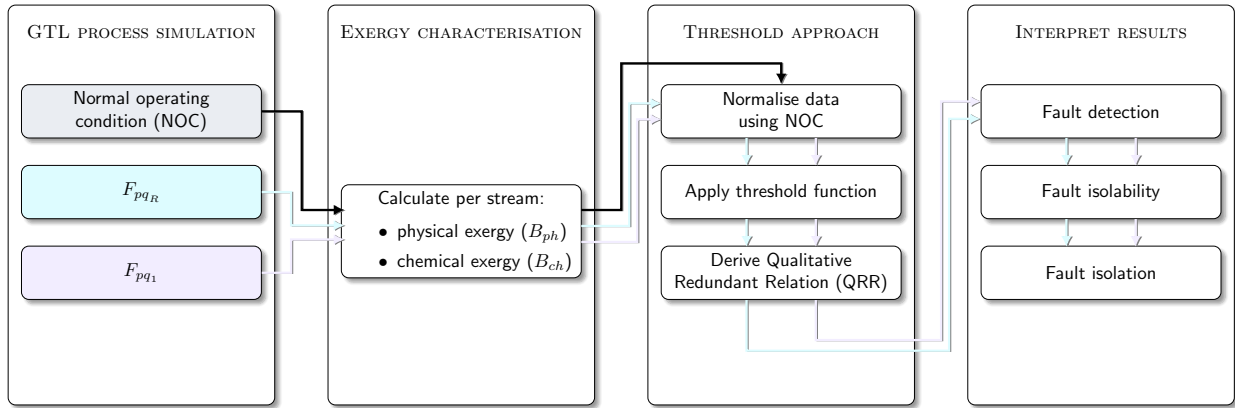


Figure 5.1: Graphical representation of the threshold approach (Approach I.A)

5.2.2 THRESHOLD APPROACH

Note that the approach was first applied to the operational conditions of dataset F_{pq_R} . To assess the approach's performance on small 3 % faults, the approach was also applied to the F_{pq_1} dataset. A step-wise breakdown of the threshold approach and how it was applied to the GTL fault data is given:

1. The exergy data per stream of the normal operating condition (NOC) and the fault datasets were prepared as discussed in Section 4.4. The exergy data are then encapsulated in a 19×2 matrix, where the rows represent the selected² GTL stream numbers and the columns the corresponding physical exergy (B_{ph}) and chemical exergy (B_{ch}) thereof. Thus, the matrix had the form:

$$F_{pq_r} = \begin{bmatrix} B_{ph(stream1)} & B_{ch(stream1)} \\ \vdots & \vdots \\ B_{ph(stream24)} & B_{ch(stream24)} \end{bmatrix} \quad (5.1)$$

2. Next, each one of the considered operational conditions was normalised with respect to the NOC exergy data (Table 4.4).
3. A simple threshold function was applied to the normalised data in order to obtain a Qualitative Redundant Relation (QRR). A QRR is a vector that indicates the qualitative variation (positive, negative, or zero) of an inspected element. In this incident, the QRR would signify the variation in exergy (both physical and chemical) between the considered operational condition and the NOC.

²To slightly reduce the number of parameters being observed, some streams such as the waste and purge streams (Streams 6, 9, 19, 22, and 23) were not included.

The threshold function used, had the form shown in Figure 5.2 and is described by:

$$y = \begin{cases} -1 & \text{if } z < (1 - \kappa) \\ 1 & \text{if } z > (1 + \kappa) \\ 0 & \text{otherwise.} \end{cases} \quad (5.2)$$

In (5.2), z represents the normalised exergy value being evaluated and y the magnitude of the resultant fault element. κ describes the threshold value of the function. In order to assign an appropriate value to κ , the deviations seen within the HYSYS[®] environment were utilised. Every time the simulation model was run, under identical operating conditions, small solver variations were found. To ensure that these simulation variations were not mistaken for faults, the variances were quantified. The details on how the threshold value κ was determined are documented in Appendix E. From these calculations, κ was found to be 0.012. Essentially, the κ -value defines a range in which any variation of the normalised exergy values are ignored (assigned a zero).

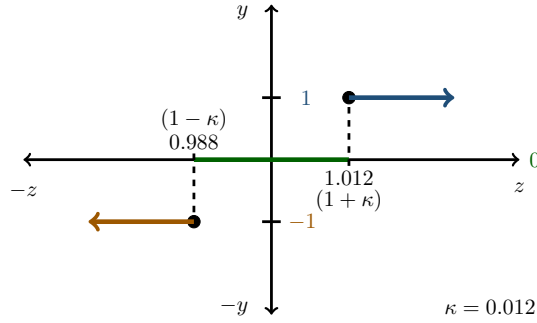


Figure 5.2: Graphical representation of the applied threshold

4. After applying the threshold function to the normalised data, a 19×2 qualitative matrix was obtained with the form

$$\mathbf{F}_{pq_r} = \begin{bmatrix} y_{B_{ph}(stream1)} & y_{B_{ch}(stream1)} \\ \vdots & \vdots \\ y_{B_{ph}(stream24)} & y_{B_{ch}(stream24)} \end{bmatrix} \quad (5.3)$$

5.2.3 ASSESSMENT METRICS

In order to evaluate how well the approach worked, a set of performance criteria was predetermined and applied. As there does not exist a single applicable set of criteria, a combination of the metrics seen in literature (summarised in Section 2.5) was used. The criteria endeavour to assess both qualitative and quantitative properties of the developed approach. Seeing as the GTL system was only considered in its steady-state, no temporal metrics were included. In the subsequent sections, the chosen metrics are discussed briefly, detailing what each one addresses and the manner in which it was achieved.

5.2.3.1 DETECTION

As already mentioned, the first task of an FDI system is to detect whether a fault is present or not. For the applied threshold approach, any non-zero qualitative matrix would indicate a fault condition. Conversely, an all-zero qualitative matrix would suggest a normal operating condition. To quantify the detection performance, a confusion matrix is drawn up [57]. The idea behind a confusion matrix is to determine the number of categorical instances achieved by an approach. Table 5.1 shows a confusion matrix and the four categories that include:

- *True negative* (TN) - the approach detected a fault-free condition and the true condition was indeed fault-free. The count is assigned to a .
- *False negative* (FN) - the approach detected a fault-free condition, but the true condition was actually faulty. The count is assigned to b .
- *False positive* (FP) - the approach detected a faulty condition, but the true condition was actually fault-free. The count is assigned to c .
- *True positive* (TP) - the approach detected a faulty condition and the true condition was indeed faulty. The count is assigned to d .

These counted instances can then be used to quantify the false positive rate (r_{FP}), the false negative rate (r_{FN}), the true positive rate (r_{TP}) and the accuracy, using the tabulated formulae. Ideally, a well-performing approach's r_{FP} and r_{FN} should be 0 % and the r_{TP} and accuracy 100 %.

Table 5.1: Confusion matrix and relevant detection rates calculations

CONFUSION MATRIX				DETECTION RATES	
		TRUE CONDITION		Rate	%
		Fault-free	Fault		
DETECTION CONDITION	Fault-free	True negative a TN	False negative b FN	r_{FP}	$\frac{c}{(a+c)} \times 100$
	Fault	False positive c FP	True positive d TP	r_{FN}	$\frac{b}{(b+d)} \times 100$
				r_{TP}	$\frac{d}{(b+d)} \times 100$
				Accuracy	$\frac{(a+d)}{(a+b+c+d)} \times 100$

5.2.3.2 ISOLABILITY

When looking at an approach's isolability performance, its ability to distinguish between faults is assessed [2, 30]. Therefore, in order for faults to be isolable, no two faults should have identical qualitative matrices. To evaluate whether the considered faults were isolable, the obtained qualitative matrices were subtracted from one another in order to obtain a distance (d_{xy}). Distance d_{xy} is computed by subtracting corresponding row (k) and column (ℓ) entries of the 19×2 qualitative matrices F_x and F_y , mathematically this is conveyed as:

$$d_{xy} = \sum_{\ell=1}^2 \sum_{k=1}^{19} |F_x(k, \ell) - F_y(k, \ell)|. \quad (5.4)$$

with $F_x(k, \ell)$ and $F_y(k, \ell)$ representing the entries in the respective fault matrices. A distance value of 0 would indicate identical qualitative entries of the two considered fault matrices; making them unisolable. The isolability performance is then expressed as a percentage of isolable conditions compared to the overall number of considered conditions.

5.2.3.3 ISOLATION

According to Severson et al. [26], isolation of a fault leads to gleaning information of the location thereof. Consequently - for this study - if the applied approach can give any feedback of where within the GTL a fault occurred, isolation of the fault is conferred. The assessment was done visually, contemplating the first non-zero qualitative matrix entry/entries and corresponding stream number. The isolation performance is also conveyed as a percentage.

5.2.3.4 SENSITIVITY

An additional desired property of an FDI approach is sensitivity. Kurtoglu et al. [57] included a metric that assesses the detection sensitivity factor which determines the relative strength of a fault when *detection* occurs. Another way to look at this would be to evaluate whether the FDI approach would be able to detect small magnitude faults. To test the sensitivity of the proposed threshold function, the results of the 3 % faults (F_{pq_1}) were investigated.

5.2.3.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

These two properties should ideally be well balanced, with relative fast algorithms and implementations and low storage requirements [2]. A qualitative assessment of both will be made once the approach is applied.

5.3 RESULTS

The threshold approach (Approach I.A) was first applied to the normal condition (Normal₁) grouped with fault dataset F_{pq_R} . Table 5.2a shows the physical and chemical exergy data of Normal₁. After normalisation, the per-stream data had the values depicted in Table 5.2b. When applying the threshold function as detailed in (5.2), the qualitative fault matrix in Table 5.2c was obtained. As anticipated, no normalised value fell beyond the threshold limits, giving an all-zero qualitative matrix. The approach was, therefore, successful in discerning a normal condition. The methodology was applied in the same manner to the remaining 12 faults of dataset F_{pq_R} and subsequently to the 3 % dataset (F_{pq_1}). Table 5.3 and Table 5.4 summarise the qualitative matrices attained for each of the considered operational conditions.

Table 5.2: Threshold approach methodology outputs per stream for normal condition Normal₁

(a) Exergy data			(b) Normalised data			(c) Qualitative matrix		
Stream no	Normal ₁		Stream no	Normal ₁		Stream no	Normal ₁	
	B_{ph}	B_{ch}		B_{ph}	B_{ch}		B_{ph}	B_{ch}
1	211991000	6811680000	1	1.000	1.000	1	0	0
2	153042000	51577400	2	1.000	1.000	2	0	0
3	42119200	17608900	3	1.000	1.000	3	0	0
4	22166700	18697700	4	1.000	1.000	4	0	0
5	853040000	5970380000	5	1.000	1.000	5	0	0
7	207502000	5920444560	7	1.000	1.000	7	0	0
8	205658000	5915080000	8	1.000	1.000	8	0	0
10	288628000	11330900000	10	0.999	0.999	10	0	0
11	300122000	11330900000	11	0.999	0.999	11	0	0
12	222295000	10651174000	12	0.998	0.999	12	0	0
13	219998000	9787070000	13	0.998	0.999	13	0	0
14	2180750	864104000	14	1.001	1.000	14	0	0
15	117153000	9734878650	15	0.997	0.999	15	0	0
16	116085000	8823660000	16	0.997	0.998	16	0	0
17	246183	1769930000	17	0.997	1.000	17	0	0
18	532825	5607760	18	0.999	1.000	18	0	0
20	92867900	7058930000	20	0.997	0.998	20	0	0
21	108638000	7058930000	21	0.998	0.998	21	0	0
24	83338600	5415780000	24	0.997	0.998	24	0	0

5.4 APPROACH PERFORMANCE EVALUATION

5.4.1 DETECTION

When visually assessing the obtained qualitative matrices in Table 5.3, it can be seen that there are no all-zero matrices. The approach could therefore detect each faulty operational condition within F_{pq_R} successfully. Subsequently, the confusion matrix - depicted in Table 5.5a - was assigned:

- *True negative* (TN) = 1 as one fault-free condition (Normal₁) was tested, and the approach was correct in its verdict.
- *True positive* (TP) = 12 as all remaining faulty operational conditions tested, were detected successfully.

Using these assignments to calculate the detection rates, excellent performance is seen. Notably the 0 % false negative rate (r_{FN}), and the detection accuracy of 100 %. The same perfect detection performance was unfortunately not seen when evaluating Table 5.4; as Approach I.A failed to detect fault F_{23_1} . The confusion matrix for dataset F_{pq_1} is given in Table 5.5b and was assigned:

- *False negative* (FN) = 1 as fault condition F_{23_1} was incorrectly seen as being fault-free.
- *True positive* (TP) = 11 as the other eleven faulty operational conditions were successfully detected.

The detection rates based on these assignments demonstrate small shortcomings in the approach's detection performance.

Table 5.3: The qualitative matrices of dataset F_{pq_R} after applying the threshold function

Stream no	F_{1q_r}						F_{2q_r}						F_{3q_r}											
	F_{116}		F_{123}		F_{137}		F_{142}		F_{213}		F_{225}		F_{236}		F_{242}		F_{317}		F_{322}		F_{335}		F_{343}	
	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}
1	1	1	-1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	-1	0	-1	-1	1	1	-1	-1
11	1	1	-1	-1	0	0	0	0	-1	0	-1	-1	0	0	0	0	0	0	-1	-1	1	1	-1	-1
12	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	-1	0	1	0	0	0	-1	-1	1	1	-1	-1
13	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	-1	0	1	1	0	0	-1	-1	1	1	-1	-1
14	1	-1	1	1	1	-1	1	-1	-1	1	1	1	-1	0	-1	-1	0	0	-1	-1	-1	-1	1	1
15	1	1	-1	-1	0	0	0	0	0	0	-1	-1	-1	0	0	1	0	0	-1	-1	1	1	-1	-1
16	1	1	-1	-1	0	0	0	0	0	0	-1	-1	-1	0	0	0	0	0	-1	-1	1	1	-1	-1
17	-1	-1	1	-1	1	1	0	0	-1	0	0	-1	-1	0	-1	0	0	0	-1	-1	-1	-1	1	1
18	-1	1	1	-1	1	1	0	0	-1	0	1	-1	-1	0	-1	0	0	0	-1	0	-1	0	1	0
20	1	1	-1	-1	0	0	0	0	0	0	-1	-1	-1	0	0	0	0	0	-1	-1	1	1	-1	-1
21	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	-1	0	-1	-1	1	1	-1	-1
24	1	1	1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	-1	0	-1	-1	1	1	-1	-1

Table 5.4: The qualitative matrices of dataset F_{pq_1} after applying the threshold function

Stream no	F_{1q_1}								F_{2q_1}								F_{3q_1}							
	F_{11_1}		F_{12_1}		F_{13_1}		F_{14_1}		F_{21_1}		F_{22_1}		F_{23_1}		F_{24_1}		F_{31_1}		F_{32_1}		F_{33_1}		F_{34_1}	
	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}
1	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	1	0	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	-1	-1	1	1	-1	-1
11	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	-1	-1	1	1	-1	-1
12	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	0	0	1	0	0	0	-1	-1	1	1	-1	-1
13	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	0	0	1	1	0	0	-1	-1	1	1	-1	-1
14	1	-1	-1	1	1	-1	1	0	-1	1	1	1	0	0	1	-1	0	0	1	0	-1	-1	1	1
15	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	1	0	0	-1	-1	1	1	-1	-1
16	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	-1	-1	1	1	-1	-1
17	-1	0	0	0	1	1	0	0	0	0	1	-1	0	0	0	0	0	0	0	0	-1	0	1	1
18	0	1	-1	-1	1	1	0	0	0	0	1	1	0	0	1	0	0	0	0	0	-1	0	1	0
20	1	1	-1	-1	0	0	0	0	0	0	1	-1	0	0	0	0	0	0	-1	-1	1	1	-1	-1
21	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	0	0	-1	-1	1	1	-1	-1
24	1	1	-1	-1	0	0	0	0	0	0	1	1	0	0	0	0	-1	0	-1	-1	1	1	-1	-1

Table 5.5: Confusion matrix when applying Approach I.A on dataset (a) F_{pq_R} and (b) F_{pq_1}

(a)				(b)			
CONFUSION MATRIX				CONFUSION MATRIX			
		TRUE CONDITION				TRUE CONDITION	
		Fault-free	Fault			Fault-free	Fault
DETECTION CONDITION	Fault-free	a TN 1	b FN 0	DETECTION CONDITION	Fault-free	a TN 0	b FN 1
	Fault	c FP 0	d TP 12		Fault	c FP 0	d TP 11
DETECTION RATES				DETECTION RATES			
Rate		%		Rate		%	
r_{FP}		0		r_{FP}		0	
r_{FN}		0		r_{FN}		8.3	
r_{TP}		100		r_{TP}		91.7	
Accuracy		100		Accuracy		91.7	

5.4.2 ISOLABILITY

For the faults to be isolable, not one qualitative matrix should be the same as another. When using (5.4) to calculate the distance d_{xy} of the qualitative matrices; any 0 entry would indicate duplicate qualitative matrices. Table 5.9 summarises the distance values for each qualitative matrix in dataset F_{pq_R} compared to all qualitative matrices within the same dataset. As expected, the diagonal values were all 0, as these were the fault matrices compared to themselves. There are no 0 entries off-diagonally, indicating unique qualitative matrices for each type of fault within dataset F_{pq_R} . The same method was applied to dataset F_{pq_1} with Table 5.10 depicting the isolability of the matrices compared to one another. Once again no off-diagonal 0 values are seen, signifying the qualitative matrices were all isolable from one another. Lastly, comparing the qualitative matrices of dataset F_{pq_R} with those of dataset F_{pq_1} - summarised in Table 5.11 - it is evident that F_{23_1} was unisolable from the normal condition. Based on these findings the approach does well to uniquely isolate the various types of faults but does not necessarily guarantee 100 % isolability, as summarised in Table 5.6.

Table 5.6: The isolability performance of Approach I.A

Fault dataset	Isolability [%]
F_{pq_R}	100.0
F_{pq_1}	96.0
Overall	97.4

5.4.3 ISOLATION

For this study, the isolation performance assesses the capability of an approach to indicate an exact fault location. By visually inspecting the qualitative matrices in Table 5.3, the exact locations of faults F_{11_6} , F_{12_3} , F_{13_7} , F_{21_3} , F_{23_6} , and F_{24_2} were found. When evaluating Table 5.4, the isolation performance decreases even further. Only being able to pinpoint the fault locations of operational conditions F_{11_1} , F_{12_1} , and F_{24_1} . A crucial aspect to take note of is the propagating effect the recycle stream has. This is particularly apparent where fault locations where beyond [Stream 11](#) but the effects are seen from [Stream 10](#). The threshold approach, therefore, does not account for these effects in any way. The quantified isolation performance of Approach I.A is summarised in Table 5.7.

Table 5.7: The isolation performance of Approach I.A

Fault dataset	Isolation [%]
F_{pq_R}	53.9
F_{pq_1}	25.0
Overall	40.0

5.4.4 SENSITIVITY

When comparing the performance metrics, summarised in Table 5.8, Approach I.A seems to have a sensitivity issue. This is evident from the fact that all of the considered performance metrics decreased for dataset F_{pq_1} . The first noticeable aspect regarding the sensitivity would be that it failed to detect any deviation within any one of the streams of F_{23_1} . If the approach were sensitive enough, the r_{FN} would be mitigated, and the detection accuracy would be 100 %. Furthermore, if the approach did detect F_{23_1} , it would have been isolable from the NOC, improving the isolability performance. Lastly, the isolation performance was influenced the most. As the threshold was not sensitive for all small deviations within the fault stream locations, some fault locations could not be determined. These findings show that sensitivity directly influences the detection, isolability and isolation performance of an FDD system.

Table 5.8: A summary of the performance metrics for Approach I.A

Fault set	r_{FP}	r_{FN}	r_{TP}	Accuracy	Isolability [%]	Isolation [%]
F_{pq_R}	0.0	0.0	100.0	100.0	100.0	53.9
F_{pq_1}	0.0	8.3	91.7	91.7	96.0	25.0

5.4.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

Little additional storage beyond the process variables already measured and stored within a process plant, would be required. To calculate the exergy of various stream constituents - which would always have a finite number of possibilities - the standard chemical exergy would be stored in a look-up table. The computational requirements of this approach would also be minimal as the technique itself only consists of two computing steps, namely the normalisation of the data and the application of the threshold values.

Table 5.9: Detection, isolability, and isolation metrics of dataset F_{pq_R} compared to itself

	Normal ₁	F_{11_6}	F_{12_3}	F_{13_7}	F_{14_2}	F_{21_3}	F_{22_5}	F_{23_6}	F_{24_2}	F_{31_7}	F_{32_2}	F_{33_5}	F_{34_3}	Detected	Isolable	Isolation
Normal ₁	0	31	32	10	2	7	23	8	8	3	23	23	23	✓	✓	✓
F_{11_6}	31	0	59	31	29	34	50	35	25	34	46	10	52	✓	✓	✓
F_{12_3}	32	59	0	34	32	31	9	30	40	29	17	53	11	✓	✓	✓
F_{13_7}	10	31	34	0	8	17	29	18	12	13	31	27	25	✓	✓	✓
F_{14_2}	2	29	32	8	0	9	23	10	8	5	23	23	23	✓	✓	✗
F_{21_3}	7	34	31	17	9	0	22	5	9	10	18	24	22	✓	✓	✓
F_{22_5}	23	50	9	29	23	22	0	21	31	20	8	44	4	✓	✓	✗
F_{23_6}	8	35	30	18	10	5	21	0	10	11	15	25	21	✓	✓	✓
F_{24_2}	8	25	40	12	8	9	31	10	0	11	23	15	31	✓	✓	✓
F_{31_7}	3	34	29	13	5	10	20	11	11	0	20	26	20	✓	✓	✗
F_{32_2}	23	46	17	31	23	18	8	15	23	20	0	36	10	✓	✓	✗
F_{33_5}	23	10	53	27	23	24	44	25	15	26	36	0	46	✓	✓	✗
F_{34_3}	23	52	11	25	23	22	4	21	31	20	10	46	0	✓	✓	✗

Table 5.10: Detection, isolability, and isolation metrics of dataset F_{pq_1} compared to itself

	F_{11_1}	F_{12_1}	F_{13_1}	F_{14_1}	F_{21_1}	F_{22_1}	F_{23_1}	F_{24_1}	F_{31_1}	F_{32_1}	F_{33_1}	F_{34_1}	Detected	Isolable	Isolation
F_{11_1}	0	58	28	28	33	43	29	24	30	46	11	50	✓	✓	✓
F_{12_1}	58	0	38	30	25	21	29	36	28	12	47	14	✓	✓	✓
F_{13_1}	28	38	0	8	13	25	9	6	10	26	25	24	✓	✓	✗
F_{14_1}	28	30	8	0	5	23	1	6	2	18	23	22	✓	✓	✗
F_{21_1}	33	25	13	5	0	22	4	11	5	19	24	21	✓	✓	✗
F_{22_1}	43	21	25	23	22	0	24	27	25	11	40	9	✓	✓	✗
F_{23_1}	29	29	9	1	4	24	0	7	1	19	22	23	✗	✓	✗
F_{24_1}	24	36	6	6	11	27	7	0	8	24	19	26	✓	✓	✓
F_{31_1}	30	28	10	2	5	25	1	8	0	18	23	22	✓	✓	✗
F_{32_1}	46	12	26	18	19	11	19	24	18	0	41	4	✓	✓	✗
F_{33_1}	11	47	25	23	24	40	22	19	23	41	0	45	✓	✓	✗
F_{34_1}	50	14	24	22	21	9	23	26	22	4	45	0	✓	✓	✗

Table 5.11: Detection, isolability, and isolation metrics of dataset F_{pq_R} compared to F_{pq_1}

	F_{11_1}	F_{12_1}	F_{13_1}	F_{14_1}	F_{21_1}	F_{22_1}	F_{23_1}	F_{24_1}	F_{31_1}	F_{32_1}	F_{33_1}	F_{34_1}	Detected	Isolable	Isolation
Normal₁	29	29	9	1	4	24	0	7	1	19	22	23	✓	✗	✓
F_{11_6}	2	58	30	30	35	43	31	26	32	48	11	52	✓	✓	✓
F_{12_3}	59	7	35	31	30	16	32	35	31	13	54	11	✓	✓	✓
F_{13_7}	29	37	1	9	14	26	10	7	11	27	26	25	✓	✓	✓
F_{14_2}	27	31	7	1	6	24	2	5	3	19	22	23	✓	✓	✗
F_{21_3}	34	24	16	8	3	23	7	14	8	20	23	22	✓	✓	✓
F_{22_5}	50	12	28	22	21	9	23	26	22	4	45	4	✓	✓	✗
F_{23_6}	35	23	17	9	6	24	8	15	9	17	24	21	✓	✓	✓
F_{24_2}	25	33	11	9	10	32	8	5	9	27	14	31	✓	✓	✓
F_{31_7}	32	26	12	4	7	23	3	10	2	16	25	20	✓	✓	✗
F_{32_2}	48	12	30	24	21	15	23	28	22	6	37	10	✓	✓	✗
F_{33_5}	12	48	26	24	25	39	23	20	24	42	1	46	✓	✓	✗
F_{34_3}	50	14	24	22	21	9	23	26	22	4	45	0	✓	✓	✗

5.5 CONCLUSION

The exergy-based threshold approach, as proposed by Marais [16], was applied to the developed GTL process which contains a recycle stream. The approach, referred to as Approach I.A, was applied to datasets F_{pq_R} and F_{pq_1} . Figure 5.3 visually summarises the cumulative performance metrics of Approach I.A. Based on the failed detection of one 3 % fault and meagre sensitivity properties; it is reasonable to state that approach's performance is not full-proof. Additionally, Approach I.A cannot indicate the exact fault locations of 15 of the 25 considered fault conditions. As such, and possibly because of the contributing propagating effects of the recycle stream, a different approach should be investigated to use as an FDI scheme for a larger-scale process such as the GTL. Chapter 6 will therefore look into the suitability of an energy-based, graph-based approach that makes use of a distance parameter (D_C -value).

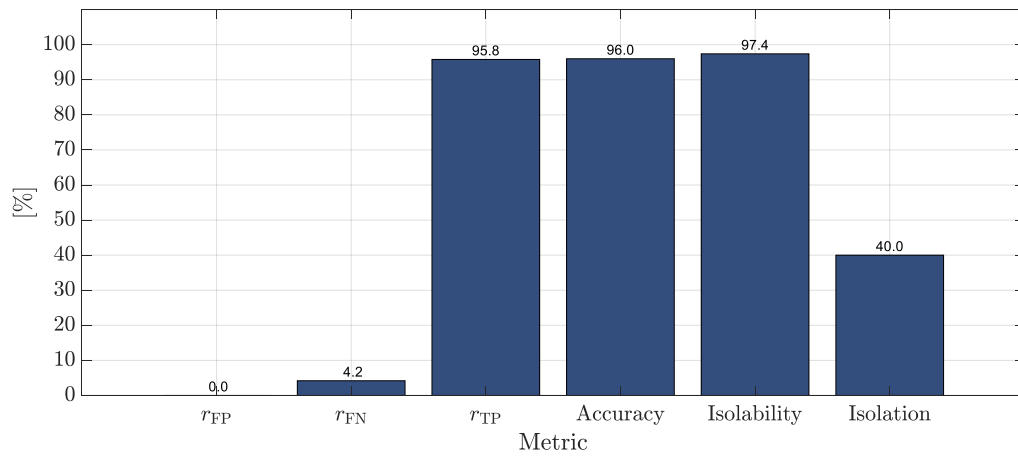


Figure 5.3: Visual depiction of the cumulative performance metrics of Approach I.A

CHAPTER 6

ENERGY-BASED FAULT DETECTION: A GRAPH MATCHING APPROACH

6.1 INTRODUCTION

The previous chapter explored the exergy-based threshold approach developed by Marais [16]. The results proved to be quite satisfactory when assessing the detection and isolability performance. The isolation performance, however, was found lacking. It should be emphasised that the exergy and structural information concepts still hold promise; the issue seems to be the fixed-threshold that was applied. Therefore, keeping with the exergy characterisation and preservation of structural information, Ould-Bouamama et al. [33] suggest that a graphical method would allow for both. Such graphical approaches would also provide different mathematical schemes of detecting and isolating considered faults. Most of the graphical approaches reviewed by [33] make use of graphs to describe system properties and relevant causal relations. Graph theory has been in use since the 1730s and became very popular in the 1930s. It is mathematical in nature, and the concepts thereof have diverse capabilities. A broad spectrum of applications is seen throughout literature, including pattern recognition, transportation and even economics. A graph essentially consists of an ordered pair $G = (V, E)$, where V is the set of vertices (also called nodes) and E the set of edges (sometimes referred to as links or arcs). Usually, vertices represent certain properties of a system, whereas the edges are used to describe the incidence relation of the vertices to themselves or other vertices within the graph, as stated by [162]. Furthermore, the graph vertices and edges can contain information. If the information is simply a name or number, the graph is called a *labelled graph*. Should additional information in the form of attributes be available, the graph is aptly named an *attributed graph*. The edges can also be either directional or have no direction related to it. From the detailed description, it is evident then why graph theory can be utilised in so many fields, notwithstanding FDD. The most suitable graphical approach for this study was chosen to be attributed graphs along with graph matching, a popular technique that quantifies the dissimilarities of compared graphs. Therefore, this chapter will determine and compare the suitability and performance of an energy-based, graph-based approach when applied to the same GTL process and energy data.

6.2 METHODOLOGY

6.2.1 QUICK OVERVIEW

The graph matching technique henceforth referred to as Approach II.B¹, is a quantitative approach which uses graph matching theoretical concepts, specifically a distance parameter (D_C -value), to describe the dissimilarities between considered graphs; an operational graph (G_o) and one of the database stored graphs (G_d). Graphs are developed using the physical structure of the GTL process and in-turn, node signature matrices are obtained by using the graph and the considered operational condition's unique energy attributes. A cost matrix (C_{od}) is then calculated by using the Heterogeneous Euclidean Overlap Metric (HEOM) that, in short, describes the distance between each row combination of the two graphs. Finally, a single distance parameter, $D_{C_{od}}$, quantifies the overall dissimilarity between the two evaluated graphs. The smaller the D_C -value, the smaller the dissimilarities are. The hypothesis is that when comparing an operational graph to the database graphs, the smallest D_C -value would indicate the most probable fault condition. A graphical representation of the approach is shown in Figure 6.1. More in-depth specifics regarding the methodology are given in the subsequent section.

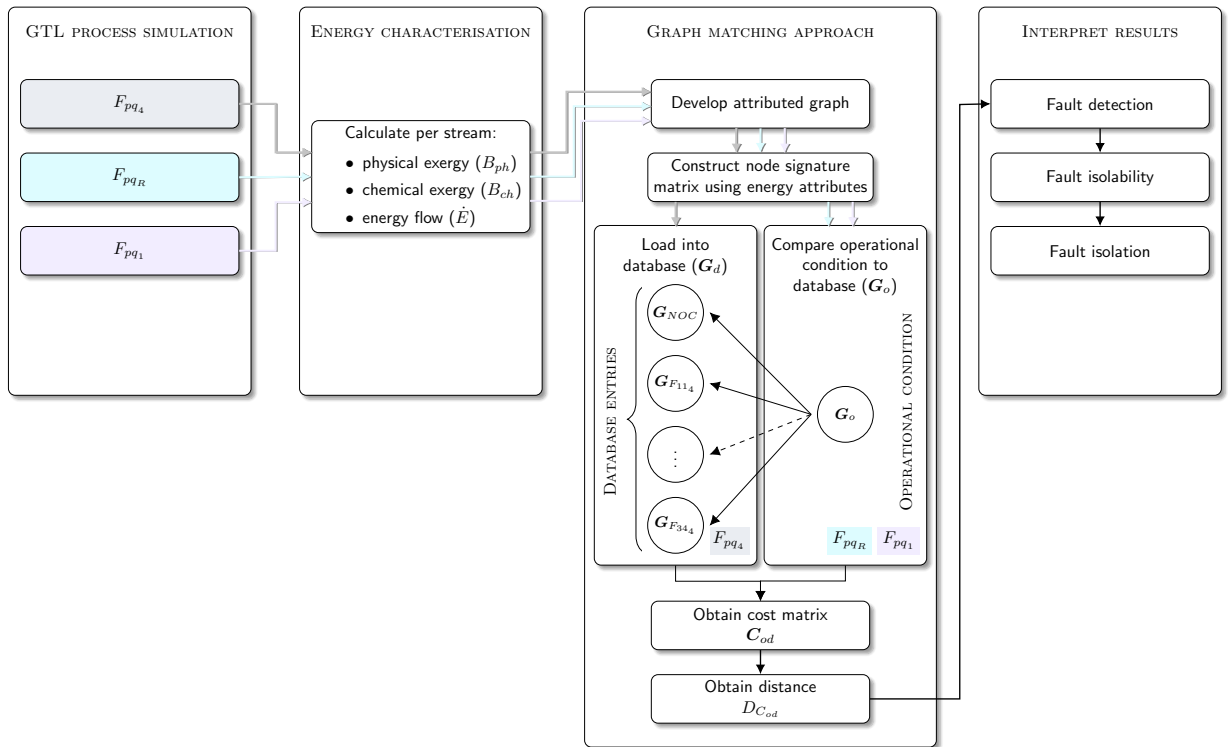


Figure 6.1: A graphical representation of the graph matching approach (Approach II.B)

6.2.2 GRAPH MATCHING APPROACH

The methodology of applying the proposed graph matching as a means to fault detection and isolation is detailed step-wise below. As mentioned, an attributed graph approach was chosen as the most suitable, as

¹Regarding the approach numbering, the A signify a qualitative approach, whereas B refers to a quantitative approach.

the graph would allow for structural information retention, and the attributes would convey numerical energy properties of the system.

1. The first important step is to develop an attributed graph of the GTL system. The graph is constructed by making use of the process flow diagram, where the:

- (a) *nodes* represent the source streams and process units.
- (b) *edges* convey the flow and connection of the units.

Table 6.1 summarises the corresponding process units and streams that were used to develop the GTL graph shown in Figure 6.2.

2. Utilising the graph and the energy data that was detailed in Chapter 4, a node signature matrix $\mathbf{G}$ is constructed with the general form:

$$\mathbf{G} = \begin{bmatrix} \Delta B_{ph_1} & \Delta B_{ch_1} & \dot{E}_{11} & \dot{E}_{12} & \dots & \dot{E}_{118} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ \Delta B_{ph_{18}} & \Delta B_{ch_{18}} & \dot{E}_{181} & \dot{E}_{182} & \dots & \dot{E}_{1818} \end{bmatrix}. \quad (6.1)$$

In (6.1), the first two columns are the *node attributes* describing the changes in physical exergy (ΔB_{ph_ι}) and chemical exergy (ΔB_{ch_ι}) over each process unit ι . The subsequent columns are *edge attributes*, specifically the energy flows ($\dot{E}_{\iota\gamma}$) between connected process units ι and γ . The number of edge attributes is dependent on the edges of each node pertaining to itself and to all other nodes. The subscripts of the edges denote the considered edges, e.g., subscript “15” would signify the edge connecting node 1 and 5. Additionally, an edge attribute is assigned to the reverse subscript notation matrix element by using the rule $\dot{E}_{\iota\gamma} = -\dot{E}_{\gamma\iota}$. Should there be no connection between two nodes, i.e., no energy flowing between two nodes; a zero is added to that matrix element. Node signature matrices were developed in this manner for each fault in datasets F_{pq_4} , F_{pq_R} , and F_{pq_1} . The general node signature matrix ($\mathbf{G}_{general}$), as well as the completed node signature matrix of the normal operating condition ($\mathbf{G}_{NOC}$), are given per illustration in (6.2) and (6.3) respectively.

3. Next, a database was constructed containing the graphs of every fault of dataset F_{pq_4} , i.e., the 10 % fault magnitudes. This was to ensure that the database contained representative characteristics of proper-sized faults. To enable simplicity of the mathematical notations, these database graphs are collectively referred to as $\mathbf{G}_d$. No graph information pertaining to the operational faults to be evaluated (F_{pq_R} and F_{pq_1}) are included in the database. These operational conditions are denoted as $\mathbf{G}_o$.

Table 6.1: The corresponding process units and streams used to construct the GTL graph

Nodes																			
Node number	①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	⑪	⑫	⑬	⑭	⑮	⑯	⑰	⑱	
Corresponding stream name	Methane	Steam	Oxygen	Carbon dioxide	-	-	-	-	-	-	-	-	-	-	Light liquids	Heavy liquids	-	-	
Corresponding process unit	-	-	-	-	ATR	Cooler 1	Separator 1	Mixer 1	Heater 1	FTR	Separator 2	Cooler 2	3 phase separator	Splitter 1	-	-	Compressor	Splitter 2	
Edges																			
Edge number	13	25	35	45	56	67	78	89	910	1011	1112	1113	1213	1314	1315	1316	1417	1718	188
Corresponding stream number	1	2	3	4	5	7	8	10	11	12	13	14	15	16	17	18	20	21	24
Corresponding stream name	Methane	Steam	Oxygen	Carbon dioxide	Syngas	Cooled syngas	Cleaned syngas	Mixed stream 1	Reactor feed	Reactor products	Gaseous products	Liquid products	Cooled reactor products	Vapour products	Light liquid products	Heavy liquid products	Recycle gas	Compressed gas	Recycle to FTR

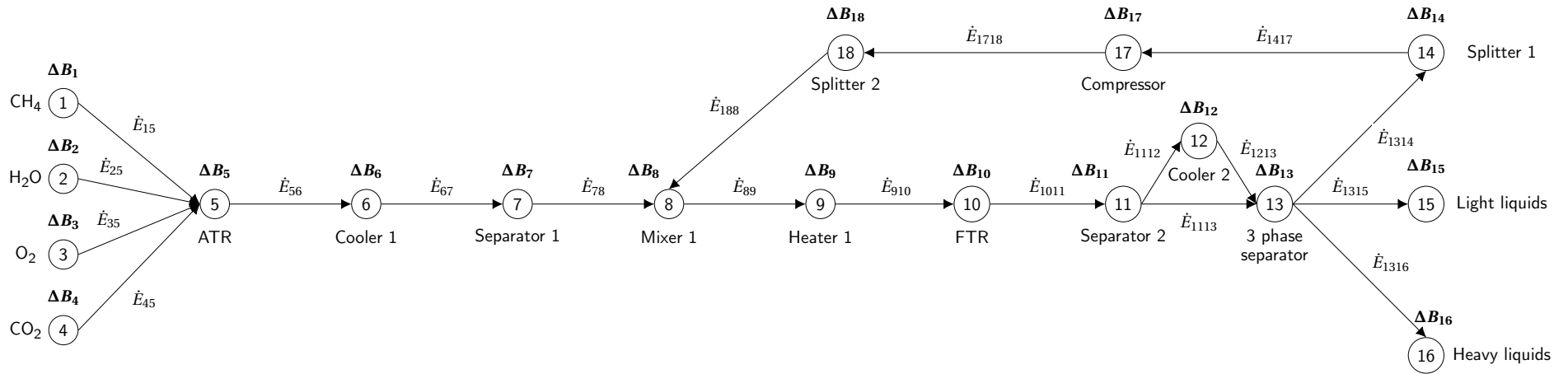


Figure 6.2: The graph of the GTL process showing the nodes, edges and energy attributes

$$G_{general} = \begin{bmatrix} \Delta B_{ph_1} & \Delta B_{ch_1} & 0 & 0 & 0 & 0 & \dot{E}_{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_2} & \Delta B_{ch_2} & 0 & 0 & 0 & 0 & \dot{E}_{25} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_3} & \Delta B_{ch_3} & 0 & 0 & 0 & 0 & \dot{E}_{35} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_4} & \Delta B_{ch_4} & 0 & 0 & 0 & 0 & \dot{E}_{45} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_5} & \Delta B_{ch_5} & -\dot{E}_{11} & -\dot{E}_{25} & -\dot{E}_{35} & -\dot{E}_{45} & 0 & \dot{E}_{56} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_6} & \Delta B_{ch_6} & 0 & 0 & 0 & 0 & -\dot{E}_{56} & 0 & \dot{E}_{67} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_7} & \Delta B_{ch_7} & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{67} & 0 & \dot{E}_{78} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_8} & \Delta B_{ch_8} & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{78} & 0 & \dot{E}_{89} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{818} \\ \Delta B_{ph_9} & \Delta B_{ch_9} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{89} & 0 & \dot{E}_{910} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_{10}} & \Delta B_{ch_{10}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{910} & 0 & \dot{E}_{1011} & 0 & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_{11}} & \Delta B_{ch_{11}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1011} & 0 & \dot{E}_{1112} & \dot{E}_{1113} & 0 & 0 & 0 & 0 \\ \Delta B_{ph_{12}} & \Delta B_{ch_{12}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1112} & 0 & \dot{E}_{1213} & 0 & 0 & 0 & 0 \\ \Delta B_{ph_{13}} & \Delta B_{ch_{13}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1113} & -\dot{E}_{1213} & 0 & \dot{E}_{1314} & \dot{E}_{1315} & \dot{E}_{1316} & 0 & 0 \\ \Delta B_{ph_{14}} & \Delta B_{ch_{14}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1314} & 0 & 0 & 0 & \dot{E}_{1417} & 0 \\ \Delta B_{ph_{15}} & \Delta B_{ch_{15}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1315} & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_{16}} & \Delta B_{ch_{16}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1316} & 0 & 0 & 0 & 0 & 0 \\ \Delta B_{ph_{17}} & \Delta B_{ch_{17}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1417} & 0 & 0 & 0 & \dot{E}_{1718} \\ \Delta B_{ph_{18}} & \Delta B_{ch_{18}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dot{E}_{188} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\dot{E}_{1718} & 0 \end{bmatrix} \quad (6.2)$$

$$G_{NOC} = \begin{bmatrix} 211991000 & 6811680000 & 0 & 0 & 0 & 0 & -329629228 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 153042000 & 51577400 & 0 & 0 & 0 & 0 & -1185124908 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 42119200 & 17608900 & 0 & 0 & 0 & 0 & 22860070 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 22166700 & 18697700 & 0 & 0 & 0 & 0 & -348963640 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 423723600 & -929184000 & 329629228 & 1185124908 & -22860070 & 348963640 & 0 & -1840857024 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -645540000 & -50466900 & 0 & 0 & 0 & 0 & 0 & 1840857024 & 0 & -3102596564 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1844000 & -4828100 & 0 & 0 & 0 & 0 & 0 & 0 & 3102596564 & 0 & -1445481368 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -356910 & 1000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1445481368 & 0 & -3032884567 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1587406162 \\ 11535300 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3032884567 & 0 & -2852584935 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -77702700 & -679252302 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2852584935 & 0 & -3978187110 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -116842 & 132626 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3978187110 & 0 & -3950984792 & -27203644 & 0 & 0 & 0 & 0 & 0 \\ -102906800 & -52179796 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3950984792 & 0 & -4378237273 & 0 & 0 & 0 & 0 & 0 \\ -2471692 & -432238 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 27203644 & 4378237273 & 0 & -2607799760 & -78084797 & -1719558002 & 0 & 0 \\ -23291650 & -1767500000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2607799760 & 0 & 0 & 0 & -2086238813 & 0 \\ -246986 & -1769064000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 78084797 & 0 & 0 & 0 & 0 & 0 \\ -533380 & -5608290 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1719558002 & 0 & 0 & 0 & 0 & 0 \\ 15682750 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2086238813 & 0 & 0 & 0 & -2066399976 \\ -25249590 & -1641182000 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1587406162 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2066399976 & 0 \end{bmatrix} \quad (6.3)$$

4. A cost matrix C_{od} is used to determine how dissimilar two graphs, G_o and G_d , are when compared to each another. To calculate this, the Heterogeneous Euclidean Overlap Metric (HEOM) proposed in the work of Jouili et al. [163] was used. Mathematically, this is conveyed by

$$C_{od(i,j)} = HEOM(i,j) = \sqrt{\sum_{\alpha=1}^A \left(\frac{|G_{oi\alpha} - G_{dj\alpha}|}{range_{\alpha}} \right)^2}, \quad (6.4)$$

eventually resulting in an $(i \times j)$ matrix. A is the number of columns of the graphs, j the number of rows in graph G_d and i the number of rows in graph G_o . To normalise the data, the $range_{\alpha}$ of each column of graph G_o is obtained and calculated by using:

$$range_{\alpha} = |max_{\alpha} - min_{\alpha}|, \quad (6.5)$$

where max_{α} is the largest numerical value and min_{α} the smallest in column α . It should be noted that by using the HEOM instead of the Euclidean distance, the following aspects are addressed:

- The HEOM can accommodate symbolic attributes, thus should some be included in the future, the proposed approach will be able to handle the additional information effortlessly.
- The Euclidean distance function does not include any normalisation, therefore, according to [164], attributes with large ranges would diminish smaller attributes' inputs.

5. In order to determine a single distance ($D_{C_{od}}$) parameter quantifying the difference between the two considered graphs G_o and G_d , the diagonal entries of the cost matrix C_{od} are summed and divided by the number of rows, i . Mathematically depicted as:

$$D_{C_{od}} = \frac{\sum_{k=1}^i C_{od(k,k)}}{i}. \quad (6.6)$$

6. The smaller the D_C -value, the smaller the dissimilarities are between the compared graphs. In other words, the likeliest match within the database is indicated by the smallest D_C -value obtained.

6.2.3 ASSESSMENT METRICS

The same assessment metrics described in Chapter 5 will be utilised when evaluating the performance of the proposed approach. The application and interpretation vary slightly; therefore, the assessment metrics are outlined below.

6.2.3.1 DETECTION

A fault is said to be detectable if the smallest D_C -value indicates any faulty condition within the database as a match. That is to say; it is undetectable should a faulty operational condition be matched with the normal NOC database graph. Based on these classifications, a confusion matrix is constructed, and relevant rates calculated.

6.2.3.2 ISOLABILITY

Isolability, which is the capability of distinguishing between two different faults, is accomplished if the proposed approach matches the operational fault to the correct database fault type. If a considered fault is matched to an incorrect fault type, the approach cannot differentiate between the two (or more), and isolability is lost. The isolability is finally expressed as a percentage of the correctly matched faults.

6.2.3.3 ISOLATION

Isolation is described as procuring information on the location of a fault. Subsequently, for this study, the definition is interpreted and assessed in the following manner. An evaluated operational fault is found to have isolation if the matched fault in the database has the same pq -assignment within the fault ID (F_{pq_r}). This is based on the fact that all fault locations of the database entries are known a priori. The isolability and isolation performance metric will inevitably have the same percentage (%) when the approach is applied in the manner discussed in this study. Nevertheless, as the two properties have different definitions, the metrics are kept separate.

6.2.3.4 SENSITIVITY

As with the threshold approach, this metric will investigate the performance when smaller 3 % faults are tested.

6.2.3.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

Once more, a qualitative assessment of the storage and computational requirements will be made.

6.3 RESULTS

The described methodology was first applied to fault dataset F_{pq_R} . The D_C -values of the normal condition (Normal₁) and each one of twelve faults as compared to the database stored graphs were recorded and is summarised in Table 6.2. The smallest D_C -value, shown highlighted, indicates the likeliest match. The same procedure was completed for fault dataset F_{pq_1} , and the results obtained given in Table 6.3.

Table 6.2: Detectability, isolability, and isolation of fault dataset F_{pq_R}

Fault ID	DATABASE FAULTS													Detected	Isolable	Isolation
	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}			
Normal ₁	0.00126	0.05044	0.06362	0.00362	0.00355	0.00260	0.09191	0.00169	0.01641	0.00142	0.05366	0.05844	0.05086	✓	✓	✓
F_{11_6}	0.05018	0.00556	0.10589	0.04919	0.04929	0.04998	0.12818	0.05010	0.04997	0.05070	0.09310	0.03835	0.08899	✓	✓	✓
F_{12_3}	0.05929	0.10922	0.01049	0.06119	0.06117	0.05901	0.07902	0.05885	0.06326	0.05881	0.05638	0.11629	0.05315	✓	✓	✓
F_{13_7}	0.00726	0.04812	0.07038	0.00458	0.00458	0.00890	0.09882	0.00822	0.02095	0.00834	0.06064	0.06090	0.05793	✓	✓	✓
F_{14_2}	0.00206	0.04839	0.06624	0.00059	0.00055	0.00405	0.09447	0.00310	0.01725	0.00332	0.05631	0.05810	0.05355	✓	✓	✓
F_{21_3}	0.00211	0.04906	0.06420	0.00434	0.00436	0.00017	0.09232	0.00175	0.01659	0.00309	0.05434	0.05684	0.05148	✓	✓	✓
F_{22_5}	0.11739	0.16831	0.08116	0.11989	0.11980	0.11688	0.01518	0.11671	0.11340	0.11684	0.06004	0.18065	0.06787	✓	✓	✓
F_{23_6}	0.00146	0.04898	0.06414	0.00382	0.00384	0.00206	0.09228	0.00037	0.01571	0.00259	0.05416	0.05683	0.05132	✓	✓	✓
F_{24_2}	0.00567	0.05032	0.06647	0.00668	0.00674	0.00767	0.09370	0.00667	0.00456	0.00661	0.05512	0.05883	0.05257	✓	✓	✓
F_{31_7}	0.00311	0.04901	0.06532	0.00547	0.00551	0.00511	0.09350	0.00407	0.01204	0.00238	0.05524	0.05662	0.05250	✓	✓	✓
F_{32_2}	0.04880	0.09490	0.05312	0.05110	0.05103	0.04844	0.05979	0.04810	0.04990	0.04827	0.01427	0.10724	0.02527	✓	✓	✓
F_{33_5}	0.05495	0.04204	0.10744	0.05627	0.05641	0.05466	0.12966	0.05489	0.05981	0.05546	0.09639	0.00674	0.09468	✓	✓	✓
F_{34_3}	0.04776	0.09958	0.05292	0.04955	0.04939	0.04811	0.06359	0.04774	0.04595	0.04715	0.02314	0.11219	0.02178	✓	✓	✓

Table 6.3: Detectability, isolability, and isolation of fault dataset F_{pq_1}

Fault ID	DATABASE FAULTS													Detected	Isolable	Isolation
	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}			
F_{11_1}	0.01370	0.03447	0.07613	0.01284	0.01296	0.01385	0.10273	0.01386	0.02045	0.01438	0.06511	0.05135	0.06207	✓	✗	✗
F_{12_1}	0.01406	0.06389	0.05221	0.01589	0.01589	0.01434	0.09127	0.01397	0.01833	0.01375	0.05321	0.07093	0.05049	✓	✗	✗
F_{13_1}	0.00778	0.04812	0.07087	0.00508	0.00510	0.00939	0.09930	0.00872	0.01341	0.00886	0.06109	0.06121	0.05837	✓	✓	✓
F_{14_1}	0.00074	0.04884	0.06515	0.00188	0.00186	0.00287	0.09338	0.00182	0.00977	0.00209	0.05523	0.05737	0.05245	✗	✗	✗
F_{21_1}	0.00079	0.04932	0.06427	0.00316	0.00315	0.00150	0.09248	0.00107	0.00976	0.00185	0.05437	0.05716	0.05156	✗	✗	✗
F_{22_1}	0.16198	0.18545	0.17694	0.16355	0.16353	0.16160	0.17489	0.16209	0.16144	0.16190	0.16488	0.18514	0.16488	✓	✗	✗
F_{23_1}	0.00012	0.04928	0.06446	0.00268	0.00265	0.00223	0.09270	0.00111	0.00941	0.00143	0.05454	0.05710	0.05177	✗	✗	✗
F_{24_1}	0.00206	0.04886	0.06506	0.00368	0.00378	0.00373	0.09293	0.00270	0.00842	0.00307	0.05450	0.05702	0.05179	✗	✗	✗
F_{31_1}	0.00071	0.04961	0.06416	0.00316	0.00313	0.00234	0.09241	0.00131	0.00938	0.00096	0.05419	0.05751	0.05139	✗	✗	✗
F_{32_1}	0.01701	0.06643	0.05301	0.01891	0.01877	0.01752	0.08161	0.01703	0.01907	0.01653	0.04127	0.07593	0.04062	✓	✗	✗
F_{33_1}	0.01544	0.03430	0.07589	0.01679	0.01695	0.01531	0.10277	0.01546	0.02252	0.01608	0.06583	0.03907	0.06356	✓	✗	✗
F_{34_1}	0.01620	0.06601	0.05405	0.01802	0.01787	0.01673	0.08251	0.01629	0.01728	0.01570	0.04341	0.07580	0.03985	✓	✗	✗

6.4 APPROACH PERFORMANCE EVALUATION

6.4.1 DETECTION

When evaluating the D_C -values, it is seen that the proposed approach correctly matched all considered operational faults within dataset F_{pq_R} to their corresponding database faults. The completed confusion matrix is shown in Table 6.4a. The approach performed quite well as there were no false negatives (FN) or false positives (FP), with both the true positives (TP) and accuracy being 100 %. However, when the approach was applied to dataset F_{pq_1} the performance drastically deteriorated. The D_C -values show poor matchings of the smaller magnitude faults. The faults that were not detected and contributed to the false negatives (FN) category, was F_{14_1} , F_{21_1} , F_{23_1} , F_{24_1} , and F_{31_1} . The confusion matrix for dataset F_{pq_1} is depicted in Table 6.4b. The false negative rate (r_{FN}) of 41.7 % and accuracy of 58.3 %, clearly indicate the poor detection performance.

Table 6.4: Confusion matrix when applying Approach II.B on dataset (a) F_{pq_R} and (b) F_{pq_1}

(a)

CONFUSION MATRIX			
		TRUE CONDITION	
		Fault-free	Fault
DETECTION CONDITION	Fault-free	<i>a</i> TN 1	<i>b</i> FN 0
	Fault	<i>c</i> FP 0	<i>d</i> TP 12

(b)

CONFUSION MATRIX				DETECTION RATES	
		TRUE CONDITION		Rate	%
		Fault-free	Fault		
DETECTION CONDITION	Fault-free	<i>a</i> TN 0	<i>b</i> FN 5	r_{FP}	0
	Fault	<i>c</i> FP 0	<i>d</i> TP 7	r_{FN}	41.7
				r_{TP}	58.3
				Accuracy	58.3

6.4.2 ISOLABILITY

When assessing the approach's ability to discern between different faults, the performance is found to be flawless when applied to F_{pq_R} . Each considered fault is correctly matched to its database counterpart. As with the detection, the performance sees a sharp decline when applied to F_{pq_1} . Only F_{13_1} was successfully isolated. The isolability performance, expressed as a percentage, is summarised in Table 6.5. The overall metric given at the bottom of the table is merely cumulative of the metrics of F_{pq_R} and F_{pq_1} .

Table 6.5: The isolability performance of Approach II.B

Fault dataset	Isolability [%]
F_{pq_R}	100.0
F_{pq_1}	8.3
Overall	56.0

6.4.3 ISOLATION

As mentioned, isolation of a fault is deemed successful if the matching of the considered fault was to a corresponding database entry of the same section and fault type; i.e. the matching shared the same pq within

their fault IDs. Analysing the results for dataset F_{pq_R} , each fault was successfully matched to its database equivalent. The same cannot be said when evaluating F_{pq_1} . The only exact location disclosed was of fault F_{13_1} . The isolation performance, taking into account correctly isolated locations, is given as a percentage in Table 6.6.

Table 6.6: The isolation performance of Approach II.B

Fault dataset	Isolation [%]
F_{pq_R}	100.0
F_{pq_1}	8.3
Overall	56.0

6.4.4 SENSITIVITY

When assessing Table 6.7, the decreased performance of Approach II.B when applied to F_{pq_1} , demonstrate the lack of sensitivity. The approach was not sensitive enough in detecting faulty conditions F_{14_1} , F_{21_1} , F_{23_1} , F_{24_1} , and F_{31_1} ; as these operational conditions were all matched to the normal condition (NOC) within the database. The isolability and isolation performance also saw a steep decline when working with 3 % faults.

Table 6.7: A summary of the performance metrics for Approach II.B

Fault set	r_{FP}	r_{FN}	r_{TP}	Accuracy	Isolability [%]	Isolation [%]
F_{pq_R}	0.0	0.0	100.0	100.0	100.0	100.0
F_{pq_1}	0.0	41.7	58.3	58.3	8.3	8.3

6.4.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

As the assessment of these requirements is qualitative, Approach I.A is used as a baseline. Seeing that a database, which stores node signature matrices, is involved; more storage is required than with Approach I.A. In order to execute (6.4) and (6.6) for each considered fault against all stored faults, the approach necessitates significantly more computations than were seen with Approach I.A.

6.5 CONCLUSION

By developing an attributed graph of the GTL process, information regarding the physical structure as well as stream composition and physical properties (described by exergy), are encapsulated. Graph theory then provides an array of methods in which to detect and isolate faults. The proposed approach assumed that the graphs of the 10 % faults were available within a database. The node signature matrix of unknown operational faults are then compared to the database faults and their dissimilarities quantified by means of first obtaining a cost matrix, followed by calculating a distance parameter D_C from the matrix. The likeliest fault is indicated by the smallest D_C -value obtained. A visual summary of the cumulative performance metrics seen throughout the chapter is shown in Figure 6.3.

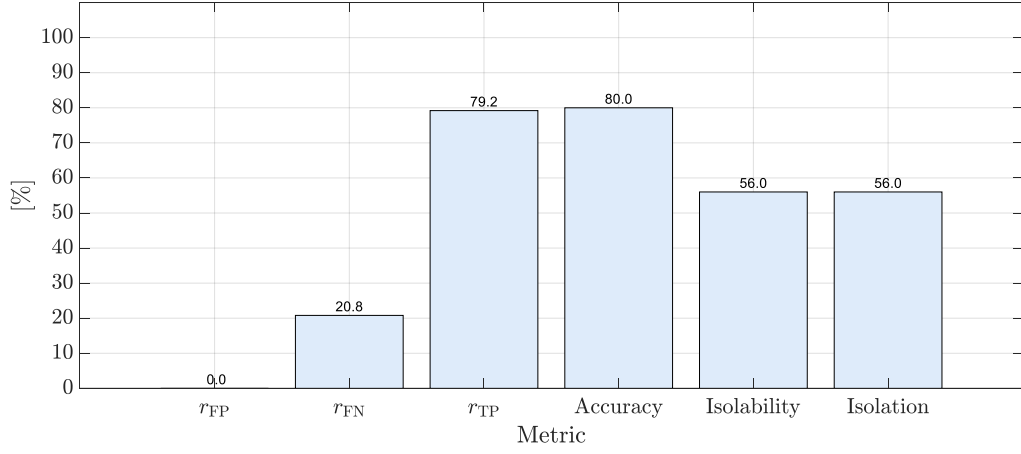


Figure 6.3: Visual depiction of the cumulative performance metrics of Approach I.A

To recapitulate, excellent performance is seen when assessing F_{pq_R} , as 100 % detection, isolability, and isolation is achieved. Regrettably, the performance decreases when presented with the smaller faults within dataset F_{pq_1} ; as many faults were seen as being normal. This signifies an issue with sensitivity, much like the threshold approach (Approach I.A) displayed. A possible reason for this could be that the chosen D_C metric discards useful information contained within the cost matrix, seeing as the 18×18 matrix is reduced to a single distance parameter. Thus, the positive performance of the F_{pq_R} dataset warrants further investigation into the benefits more degrees of freedom would bring about, particularly considering sensitivity. Chapter 7 will look into employing eigendecomposition of the cost matrix and the FDI performance thereof. An overall comparison can then be drawn between the investigated approaches.

CHAPTER 7

EXERGY-BASED FAULT DETECTION: EIGENDECOMPOSITION APPROACH

7.1 INTRODUCTION

The energy-based, graph-based approach documented in Chapter 6 showed auspicious FDI performance. Some sensitivity issues were encountered, however. A possible reason for this could be that useful information is lost when the 18×18 cost matrices are reduced to a single distance (D_C) parameter. It is, therefore, prudent to investigate the effects more matching parameters will have on the performance. Novel work done by van Graan [165], Uren et al. [19, 166], and Naser [18] promote the utilisation of eigendecomposition to diagnose faults within a system. The eigendecomposition entails the analysis of the cost matrix's eigenvalues and eigenvectors. According to [167], and the main rationale behind why these researchers [18, 19, 165, 166] championed the technique, is that eigenvalues and eigenvectors give useful descriptions of a matrix's structure and characteristics. This chapter will examine the suitability of utilising eigenvalues to achieve FDI. The methodology is first outlined before presenting the results that were obtained. The interpretations are discussed, and finally, the comparison between the explored approaches are reviewed.

7.2 METHODOLOGY

7.2.1 QUICK OVERVIEW

The approaches investigated in this chapter are based on similar fundamentals as the graph matching approach discussed in Chapter 6. They follow the same methodology up to a certain point, with the main difference being the matching mechanism. The distance (D_C) parameter is not calculated or utilised, rather the cost matrices' 18×1 eigenvalues are used. A graphical representation of the approach is shown in Figure 7.1. The first approach considers the eigenvalues qualitatively (Approach III.A), whereas the second approach makes use of exact eigenvalue-quantities (Approach III.B). A database is once again constructed, containing the node signature matrices of the NOC and F_{pq_4} fault conditions (collectively referred to as G_d). The

considered operational condition (G_o) is firstly compared to itself, i.e., the cost matrix C_{oo} is obtained. Next, the operational graph is compared to all the database stored graphs to calculate those cost matrices (C_{od}). Subsequently, the eigenvalues [$\bar{\lambda}_{oo}$ and $\bar{\lambda}_{od}$] of each the mentioned cost matrices are computed. The eigenvalues are then utilised to achieve matchings of the operational condition [$\bar{\lambda}_{oo}$] to all database entries' [$\bar{\lambda}_{oNOC}, \bar{\lambda}_{oF_{114}}, \dots, \bar{\lambda}_{oF_{344}}$], the specifics of which is given in Section 7.2.2.1 and Section 7.2.2.2. The choice of comparing the operational condition to itself and then to other database entries is based on the manner in which the graph matching using a distance parameter was set up. Note that previously and in order to match an operational condition to a database entry, the cost matrix of the two graphs were simply calculated and used to obtain the D_C -value; of which the quantitative value was a direct indication of the dissimilarity. For this approach, the eigenvalues of the cost matrices are used as a matching mechanism, and as such, it is necessary to compare the operational condition to itself to obtain an applicable cost matrix to match to database entries.

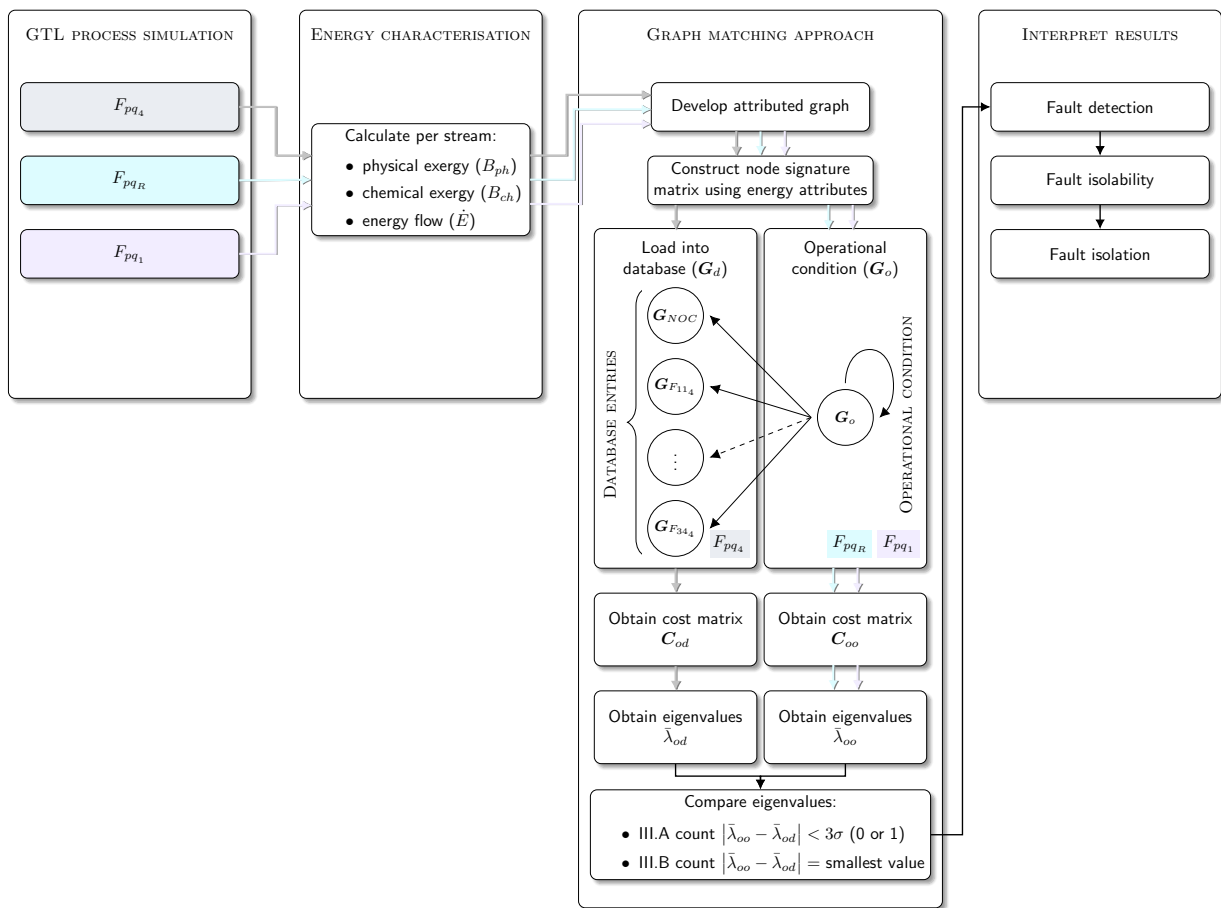


Figure 7.1: A graphical representation of the eigendecomposition approach (Approach III)

7.2.2 EIGENDECOMPOSITION

1. The same attributed graph, as developed in Chapter 6 (Figure 6.2), was utilised.
2. The attributes remained the same, recapitulation thereof being that:

(a) *node attributes* are the change in exergy (ΔB) over the process unit

- (b) *edge attributes* are the energy flows ($E_{\iota\gamma}$) between connected process units ι and γ
3. The node signature matrix was constructed in the same manner as detailed in Section 6.2.2.
 4. Once again a database was developed, comprising of the NOC and the 12 F_{pq_4} faults' node signature matrices.
 5. The cost matrix of the operational condition (G_o) compared to itself is firstly calculated making use of the Heterogeneous Euclidean Overlap Metric (HEOM):

$$C_{oo(i,j)} = HEOM(i,j) = \sqrt{\sum_{a=1}^A \frac{|G_{oia} - G_{oja}|}{range_a}} \quad (7.1)$$

6. Next, the cost matrix of the operational matrix (G_o) compared to all database entries (G_d) are calculated using the HEOM (C_{od}).
7. The 18×1 eigenvalues $\bar{\lambda} = [\lambda_{(1)}, \dots, \lambda_{(18)}]$ of each of the cost matrices are obtained by making use of the MATLAB[®] function `D = eig(C, 'vector')`.
8. Lastly, the comparison of the eigenvalues are then done qualitatively (Approach III.A) and quantitatively (Approach III.B); in order to determine whether one approach would achieve better matchings than the other.

7.2.2.1 QUALITATIVE

- (a) A well known statistical hypothesis test for determining outliers within a dataset is the 68–95–99.7 rule, alternatively called the 3σ -rule. The rule states that any observation beyond three times the standard deviation of the data would most likely be improbable and as such, an outlier. It should be noted that the 3σ -rule is only used to quantify a threshold and is not employed as a standalone test for this approach.
- (b) The qualitative method entails assigning a 0 or 1, based on the eigenvalue comparison being more or less than three times the standard deviation¹, described mathematically as:

$$y = \begin{cases} 0 & \text{if } |\bar{\lambda}_{oo} - \bar{\lambda}_{od}| < 3\sigma \\ 1 & \text{otherwise.} \end{cases} \quad (7.2)$$

- (c) These assignments are then summed per database comparison, with the smallest summed-value indicating the database fault that is the most similar to the considered operational condition. As an illustrative example, the qualitative assignments of the comparison between operational condition F_{13_7} 's eigenvalues and the database entries' eigenvalues are depicted in Table 7.1. Seeing as database fault F_{13_4} had the least number of deviations beyond 3σ , the approach was correct in its matching.

¹See Appendix F for the standard deviation calculations

7.2.2.2 QUANTITATIVE

- (a) For the quantitative approach, the smallest **numerical** difference calculated between compared eigenvalues were used to determine the matches.
- (b) Operational condition F_{21_3} is shown in Table 7.2 as an example. Looking at the first eigenvalue ($\lambda_{(1)}$), the smallest numerical difference in compared eigenvalues is found to be that of $|\lambda_{oo_{(1)}} - \lambda_{oF_{23_4(1)}}|$. As such, the eigenvalue assignment goes to the $F_{23_4(1)}$ database condition. Similarly, the smallest numerical value seen for the third eigenvalue, is $|\lambda_{oo_{(3)}} - \lambda_{oF_{21_4(3)}}|$. The approach continues until the smallest numerical value assignment is completed for all 18 eigenvalues. The condition with the most number of smallest values is deemed the designated match. For this specific example, the likeliest match would be database entry F_{21_4} as it has 15 assignments.

Table 7.1: Qualitative assignments using fault condition F_{13_7} as operational example

$\bar{\lambda}$	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}
$\lambda_{(1)}$	0	1	1	0	0	0	1	0	1	0	1	1	1
$\lambda_{(2)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\lambda_{(3)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\lambda_{(4)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\lambda_{(5)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\lambda_{(6)}$	0	1	1	0	0	0	1	0	1	0	1	1	1
$\lambda_{(7)}$	0	1	1	0	0	1	1	1	1	1	1	1	1
$\lambda_{(8)}$	0	1	1	0	0	1	1	1	1	1	1	1	1
$\lambda_{(9)}$	1	0	1	0	1	1	1	1	1	1	1	1	1
$\lambda_{(10)}$	0	1	1	0	0	0	1	1	1	1	1	1	1
$\lambda_{(11)}$	0	1	1	0	0	0	1	1	1	1	1	1	1
$\lambda_{(12)}$	1	1	1	0	0	1	1	1	1	1	1	1	1
$\lambda_{(13)}$	0	0	1	0	0	0	1	1	1	1	1	1	1
$\lambda_{(14)}$	0	0	1	0	0	1	1	0	1	1	1	1	1
$\lambda_{(15)}$	1	0	1	1	1	1	1	1	1	1	1	0	1
$\lambda_{(16)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\lambda_{(17)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
$\lambda_{(18)}$	1	1	1	1	1	1	1	1	1	1	1	1	1
	10	14	18	8	9	13	18	15	18	16	18	17	18

Table 7.2: Example of quantitative differences of fault condition F_{213}

$\bar{\lambda}$	NOC	F_{114}	F_{124}	F_{134}	F_{144}	F_{214}	F_{224}	F_{234}	F_{244}	F_{314}	F_{324}	F_{334}	F_{344}
$\lambda_{(1)}$	0.00861	0.55332	0.50920	0.01487	0.01428	0.00077	0.94471	0.00008	0.05763	0.00361	0.52152	0.64285	0.50411
$\lambda_{(2)}$	0.00000	0.00010	0.00033	0.00001	0.00001	0.00000	0.00210	0.00000	0.00036	0.00000	0.00037	0.00007	0.00104
$\lambda_{(3)}$	0.00017	0.00818	0.01357	0.00029	0.00029	0.00004	0.06262	0.00011	0.02068	0.00021	0.02516	0.00561	0.04526
$\lambda_{(4)}$	0.00046	0.00579	0.03575	0.00108	0.00107	0.00003	0.06116	0.00028	0.01090	0.00050	0.01454	0.00457	0.02241
$\lambda_{(5)}$	0.00024	0.00285	0.00454	0.00114	0.00115	0.00001	0.04193	0.00013	0.00051	0.00028	0.03778	0.00018	0.00564
$\lambda_{(6)}$	0.00121	0.03212	0.08950	0.00214	0.00205	0.00004	0.11549	0.00110	0.00834	0.00206	0.10932	0.05266	0.09634
$\lambda_{(7)}$	0.00363	0.02871	0.11039	0.00399	0.00407	0.00036	0.19732	0.00381	0.02121	0.00658	0.11429	0.03948	0.09655
$\lambda_{(8)}$	0.00224	0.02591	0.11612	0.00308	0.00308	0.00018	0.19971	0.00263	0.01462	0.00416	0.18404	0.03975	0.11227
$\lambda_{(9)}$	0.00275	0.00233	0.08571	0.00453	0.00461	0.00006	0.14374	0.00285	0.01341	0.00485	0.10713	0.00870	0.08105
$\lambda_{(10)}$	0.00228	0.01813	0.08509	0.00264	0.00273	0.00011	0.17722	0.00303	0.01573	0.00420	0.14206	0.02209	0.09446
$\lambda_{(11)}$	0.00295	0.01757	0.12037	0.00333	0.00335	0.00010	0.17957	0.00320	0.02517	0.00453	0.09907	0.01405	0.09591
$\lambda_{(12)}$	0.00289	0.01492	0.15888	0.00401	0.00372	0.00019	0.24157	0.00286	0.01839	0.00490	0.17387	0.02153	0.12801
$\lambda_{(13)}$	0.00159	0.00639	0.12795	0.00261	0.00191	0.00000	0.19046	0.00214	0.00765	0.00414	0.10731	0.00963	0.09827
$\lambda_{(14)}$	0.00450	0.00136	0.13700	0.00454	0.00451	0.00037	0.20405	0.00465	0.00939	0.00909	0.15055	0.00615	0.12350
$\lambda_{(15)}$	0.00410	0.00328	0.17323	0.00481	0.00463	0.00027	0.18820	0.00394	0.03314	0.00562	0.09080	0.00275	0.10005
$\lambda_{(16)}$	0.00053	0.02215	0.06462	0.01522	0.01519	0.00005	0.12138	0.00014	0.01197	0.00115	0.03261	0.00620	0.04839
$\lambda_{(17)}$	0.00152	0.06699	0.22204	0.00550	0.00635	0.00019	0.42537	0.00071	0.10511	0.00202	0.14185	0.05533	0.25034
$\lambda_{(18)}$	0.00167	0.07294	0.11965	0.00433	0.00548	0.00023	0.05459	0.00019	0.03959	0.00222	0.03114	0.10895	0.03120
	0	0	0	0	0	15	0	3	0	0	0	0	0

7.2.3 ASSESSMENT METRICS

7.2.3.1 DETECTION

For both the qualitative (III.A) and quantitative (III.B) approach, a fault is said to be detectable if the matched database entry can successfully detail whether the considered operational condition was faulty or normal. Based on these classifications, a confusion matrix is once again constructed, and the relevant rates calculated.

7.2.3.2 ISOLABILITY

Isolability, which is the capability of distinguishing between two different faults, is accomplished if the proposed approach matches the operational fault to the correct database fault type. If a considered fault is matched to an incorrect fault type, the approach cannot differentiate between the two (or more), and isolability is lost. The isolability is finally expressed as a percentage of the correctly matched faults.

7.2.3.3 ISOLATION

Isolation is described as observing information on the possible location of a fault. Subsequently, an evaluated fault is found to have isolation if the matched fault in the database has the same pq assignment within the fault ID F_{pq_r} .

7.2.3.4 SENSITIVITY

As with the previous approaches, this metric will investigate the performance when smaller 3 % faults are evaluated.

7.2.3.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

Qualitative assessments of the storage and computational requirements will be made.

7.3 RESULTS

The methodology of Approach III.A was applied first, to both datasets F_{pq_R} and F_{pq_1} . The smallest number of qualitative assignments when comparing the operational conditions to the database stored conditions is shown highlighted in Table 7.3 and Table 7.4. Table 7.5 and Table 7.6 tabulates the number of quantitative matches when Approach III.B is applied to F_{pq_R} and F_{pq_1} , respectively.

7.3.1 APPROACH III.A

Table 7.3: Detectability, isolability and isolation of fault dataset F_{pqR} when applying Approach III.A

Fault ID	DATABASE FAULTS													Detected	Isolable	Isolation
	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}			
Normal ₁	0	15	18	5	5	4	18	4	17	0	18	16	18	✗	✗	✗
F_{11_6}	18	17	18	18	18	18	17	18	18	18	18	18	18	✓	✗	✗
F_{12_3}	17	16	18	17	17	16	18	17	17	17	17	17	18	✓	✗	✗
F_{13_7}	10	14	18	8	9	13	18	15	18	16	18	17	18	✓	✓	✓
F_{14_2}	5	15	18	1	1	9	18	6	17	5	18	16	18	✓	✗	✗
F_{21_3}	2	15	18	7	7	0	18	1	18	6	18	15	18	✓	✓	✓
F_{22_5}	17	17	18	17	17	17	14	17	17	17	18	18	18	✓	✓	✓
F_{23_6}	2	15	18	5	6	3	18	0	17	4	18	16	18	✓	✓	✓
F_{24_2}	14	15	18	13	13	15	18	15	11	14	18	16	18	✓	✓	✓
F_{31_7}	5	15	18	11	11	12	18	12	18	4	18	17	18	✓	✓	✓
F_{32_2}	16	17	17	16	16	16	17	16	16	16	18	17	18	✗	✗	✗
F_{33_5}	18	18	18	18	18	18	17	18	18	18	18	17	18	✓	✗	✗
F_{34_3}	16	15	18	16	16	16	18	16	16	16	18	18	14	✓	✓	✓

Table 7.4: Detectability, isolability and isolation of fault dataset F_{pq_1} when applying Approach III.A

Fault ID	DATABASE FAULTS													Detected	Isolable	Isolation
	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}			
F_{11_1}	18	14	17	18	18	18	18	18	18	18	17	17	18	✓	✓	✓
F_{12_1}	12	16	18	14	14	13	18	13	15	12	18	16	18	✗	✗	✗
F_{13_1}	11	14	18	9	9	15	18	15	16	16	18	17	18	✓	✗	✗
F_{14_1}	5	15	18	4	4	5	18	5	17	5	18	16	18	✓	✗	✗
F_{21_1}	0	15	18	5	5	4	18	3	17	0	18	16	18	✗	✗	✗
F_{22_1}	18	18	18	18	18	18	18	18	18	18	18	18	18	✗	✗	✗
F_{23_1}	0	15	18	4	5	5	18	4	16	0	18	16	18	✗	✗	✗
F_{24_1}	1	15	18	4	5	10	18	6	17	6	18	16	18	✗	✗	✗
F_{31_1}	0	15	18	5	5	5	18	4	16	0	18	16	18	✗	✗	✗
F_{32_1}	12	16	18	15	14	12	18	12	17	12	18	16	18	✗	✗	✗
F_{33_1}	18	15	17	17	18	18	18	18	17	18	18	16	18	✓	✗	✗
F_{34_1}	12	15	18	13	13	12	18	12	16	12	18	17	18	✗	✗	✗

7.3.2 APPROACH III.B

Table 7.5: Detectability, isolability and isolation of fault dataset F_{pqR} when applying Approach III.B

Fault ID	DATABASE FAULTS													Detected	Isolable	Isolation
	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}			
Normal ₁	8	0	0	0	0	2	0	2	0	6	0	0	0	✓	✓	✓
F_{11_6}	0	17	0	0	0	0	1	0	0	0	0	0	0	✓	✓	✓
F_{12_3}	1	1	9	0	0	3	0	0	0	0	2	2	0	✓	✓	✓
F_{13_7}	0	3	0	9	5	0	0	1	0	0	0	0	0	✓	✓	✓
F_{14_2}	1	0	0	3	13	1	0	0	0	0	0	0	0	✓	✓	✓
F_{21_3}	0	0	0	0	0	15	0	3	0	0	0	0	0	✓	✓	✓
F_{22_5}	0	0	1	0	0	1	15	0	1	0	0	0	0	✓	✓	✓
F_{23_6}	1	0	0	0	0	1	0	16	0	0	0	0	0	✓	✓	✓
F_{24_2}	1	1	0	5	2	0	0	1	7	1	0	0	0	✓	✓	✓
F_{31_7}	11	2	0	0	0	0	0	0	0	5	0	0	0	✗	✗	✗
F_{32_2}	0	1	1	0	0	2	1	3	2	1	5	0	2	✓	✓	✓
F_{33_5}	0	1	0	0	0	0	1	0	0	0	0	16	0	✓	✓	✓
F_{34_3}	0	1	1	0	1	0	0	0	7	0	1	0	7	✓	✗	✗

Table 7.6: Detectability, isolability and isolation of fault dataset F_{pq_1} when applying Approach III.B

Fault ID	DATABASE FAULTS													Detected	Isolable	Isolation
	NOC	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}	F_{34_4}			
F_{11_1}	0	10	0	1	1	1	0	0	2	0	1	2	0	✓	✓	✓
F_{12_1}	2	0	0	1	0	4	0	4	2	4	0	1	0	✓	✗	✗
F_{13_1}	0	4	0	8	1	0	0	1	4	0	0	0	0	✓	✓	✓
F_{14_1}	11	0	0	1	5	0	0	0	0	1	0	0	0	✗	✗	✗
F_{21_1}	14	0	0	0	0	1	0	1	0	2	0	0	0	✗	✗	✗
F_{22_1}	0	0	1	0	0	0	2	0	1	0	12	2	0	✓	✗	✗
F_{23_1}	17	0	0	0	0	0	0	0	0	1	0	0	0	✗	✗	✗
F_{24_1}	7	2	0	4	2	0	0	1	0	2	0	0	0	✗	✗	✗
F_{31_1}	14	0	0	0	0	0	0	2	0	2	0	0	0	✗	✗	✗
F_{32_1}	2	1	0	2	2	1	0	2	1	6	1	0	0	✓	✗	✗
F_{33_1}	0	7	1	1	1	2	0	0	2	0	0	4	0	✓	✗	✗
F_{34_1}	2	1	0	0	2	2	0	0	4	6	1	0	0	✓	✗	✗

7.4 APPROACH PERFORMANCE EVALUATION

This section will discuss the various performance metrics, in its entirety, firstly for Approach III.A and then for Approach III.B. As one of the outcomes is to determine whether qualitative or quantitative would perform better, the comparison between the two approaches' performance is highlighted.

7.4.1 APPROACH III.A

7.4.1.1 DETECTION

When evaluating Table 7.3, the first concern is that the approach could not distinguish whether the normal operational condition (Normal₁) considered was normal or faulty; as depicted by the matches of database entries NOC and F_{314} . For this reason, the condition contributed to the false positive (FP) within the confusion matrix, given in Table 7.7a. Similarly, faulty operational condition F_{322} was indistinguishable from being normal or faulty. As such, a 1 was assigned to the false negative (FN). The remaining 11 faults were all found to be faulty, i.e., true positives (TP). Therefore, an overall detection accuracy of 84.6 % is obtained for dataset F_{pq_R} . When evaluating the qualitative results attained for dataset F_{pq_1} (Table 7.4), the false negative (FN) rate increases, as faulty operational conditions F_{121} , F_{211} , F_{221} , F_{231} , F_{241} , F_{311} , F_{321} , and F_{341} were all indiscernible from the NOC. Only 4 faults were successfully detected, with the detection accuracy dropping to 33.3 %.

Table 7.7: Confusion matrix when applying Approach III.A to dataset (a) F_{pq_R} and (b) F_{pq_1}

(a)				(b)			
CONFUSION MATRIX				CONFUSION MATRIX			
		TRUE CONDITION				TRUE CONDITION	
		Fault-free	Fault			Fault-free	Fault
DETECTION CONDITION	Fault-free	<i>a</i> TN 0	<i>b</i> FN 1	DETECTION CONDITION	Fault-free	<i>a</i> TN 0	<i>b</i> FN 8
	Fault	<i>c</i> FP 1	<i>d</i> TP 11		Fault	<i>c</i> FP 0	<i>d</i> TP 4
DETECTION RATES				DETECTION RATES			
		Rate	%			Rate	%
r_{FP}		100		r_{FP}		0	
r_{FN}		8.3		r_{FN}		66.7	
r_{TP}		91.7		r_{TP}		33.3	
Accuracy		84.6		Accuracy		33.3	

7.4.1.2 ISOLABILITY

Observing the qualitative matches summarised in Table 7.3, 7 faults (F_{137} , F_{213} , F_{225} , F_{236} , F_{242} , F_{317} , and F_{343}) were found to be isolable. Thus, the isolability performance for dataset F_{pq_R} was 53.8 %. The performance drastically decreases to only 8.3 % when assessing Table 7.4, as only F_{111} was isolable.

Table 7.8: The isolability performance of Approach III.A

Fault dataset	Isolability [%]
F_{pq_R}	53.8
F_{pq_1}	8.3
Overall	32.0

7.4.1.3 ISOLATION

For dataset F_{pq_R} , the approach could successfully indicate the faulty location of faults F_{13_7} , F_{21_3} , F_{22_5} , F_{23_6} , F_{24_2} , F_{31_7} , and F_{34_3} . When considering dataset F_{pq_1} , only one exact location - F_{11_1} - was indicated. The isolation performance for approach III.A is tabulated in Table 7.9.

Table 7.9: The isolation performance of Approach III.A

Fault dataset	Isolation [%]
F_{pq_R}	53.8
F_{pq_1}	8.3
Overall	32.0

7.4.1.4 SENSITIVITY

Approach III.A was investigated as an attempt to improve the graph matching sensitivity. Unfortunately, the matchings did not benefit from the additional number of observed parameters (eigenvalues). This is evident when comparing the performance - specifically the detection accuracy - of Approach III.A and Approach II.B applied to dataset F_{pq_1} , summarised in Table 7.10.

Table 7.10: Comparison of the performance metrics of dataset F_{pq_1} for Approach II.B and Approach III.A

Approach	r_{FP}	r_{FN}	r_{TP}	Accuracy	Isolability [%]	Isolation [%]
II.B	0.0	41.7	58.3	58.3	8.3	8.3
III.A	0.0	66.7	33.3	33.3	8.3	8.3

7.4.1.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

Once again a database storing node signature matrices is required, with the addition of having to store the calculated eigenvalues. Thus, slightly more storage is needed compared to that of Approach II.B. For this approach, the HEOM was calculated as it was done for Approach II.B. Instead of determining the distance parameter from the cost matrices, the eigenvalues were computed. This requires marginally more computational power. Furthermore, the qualitative assignments made and the eventual matchings summed, both contribute to auxiliary computational requirements.

7.4.1.6 SUMMARISING REMARKS

When assessing the individual dataset performance metrics of Approach III.A given in Table 7.11 it is seen that it is the first of the examined approaches to have 100 % false positive rate (r_{FP}), i.e., that could not discern the operational normal (Normal_1) from a faulty condition. Moreover, it is seen that the performance of the approach when applied to the 3 % faults, dramatically declines. The cumulative detection accuracy is also deemed poor at only 60 % (see Figure 7.2).

Table 7.11: A summary of the performance metrics for Approach III.A

Fault set	r_{FP}	r_{FN}	r_{TP}	Accuracy	Isolability [%]	Isolation [%]
F_{pq_R}	100.0	8.3	91.7	84.6	53.8	53.8
F_{pq_1}	0.0	66.7	33.3	33.3	8.3	8.3

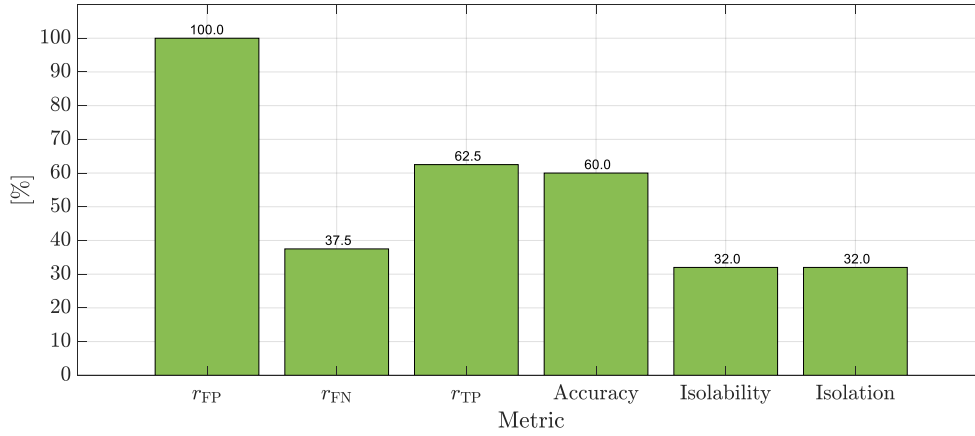


Figure 7.2: Visual depiction of the overall performance metrics of Approach III.A

7.4.2 APPROACH III.B

7.4.2.1 DETECTION

Assessing the results that were obtained, shown in Table 7.5, when applying Approach III.B to dataset F_{pq_R} ; the approach is seen to give one false negative (FN). All other considered operational conditions were correctly classified, achieving a detection accuracy of 92.3 %. Looking at dataset F_{pq_1} , 5 faults (F_{14_1} , F_{21_1} , F_{23_1} , F_{24_1} , and F_{31_1}) were not detected. This decreased the detection accuracy to only 58.3 %.

Table 7.12: Confusion matrix when applying Approach III.B to dataset (a) F_{pq_R} and (b) F_{pq_1}

(a)				(b)			
CONFUSION MATRIX				DETECTION RATES		CONFUSION MATRIX	
DETECTION CONDITION		TRUE CONDITION		Rate	%	DETECTION CONDITION	
		Fault-free	Fault				
	Fault-free	a TN 1	b FN 1	r_{FP}	0		Fault-free
	Fault	c FP 0	d TP 11	r_{FN}	8.3		Fault
				r_{TP}	91.7		
				Accuracy	92.3		

The cumulative detection rates and accuracy of Approach III.A and Approach III.B, considering the detection performance of both fault datasets, are tabulated in Table 7.13. When examining the metrics, it is evident that a **quantitative** route improves the detection performance, with the following supporting the determination:

- The false positive (FP) of Approach III.A was mitigated.
- The false negative rate (r_{FN}) is slightly improved from 37.5 % to 25.0 %.

- A greater number of faults were detected, bringing the true positive rate (r_{TP}) up to 75 %.
- The detection accuracy increased from 60.0 % to 76.0 %.

Table 7.13: Comparison of the overall detection metrics for Approach III.A and Approach III.B

Approach	r_{FP}	r_{FN}	r_{TP}	Accuracy
III.A	100.0	37.5	62.5	60.0
III.B	0.0	25.0	75.0	76.0

7.4.2.2 ISOLABILITY

When considering the isolability of Approach III.B, only faults F_{317} and F_{343} of dataset F_{pq_R} were not isolable. Once again a sharp decrease is seen when evaluating the smaller faults of F_{pq_1} ; with only two faults, F_{111} and F_{131} , being isolable. Reiteratively, the isolation performance summarised in Table 7.14, shows that the quantitative approach (III.B) performs better than the qualitative approach (III.A).

Table 7.14: Comparison of isolability performance of Approach III.A and Approach III.B

Approach	Isolability [%]		
	F_{pq_R}	F_{pq_1}	Cumulative
III.A	53.8	8.3	32.0
III.B	84.6	16.7	52.0

7.4.2.3 ISOLATION

As expected, the same faults (F_{317} and F_{343}) that were found to be unisolable, gave no information on the relevant fault-location. For dataset F_{pq_1} only two fault locations were indicated, (F_{111} and F_{131}). Table 7.15 shows once more that Approach III.B delivers better performance than III.A.

Table 7.15: Comparison of isolation performance of Approach III.A and Approach III.B

Approach	Isolation [%]		
	F_{pq_R}	F_{pq_1}	Overall
III.A	53.8	8.3	32.0
III.B	84.6	16.7	52.0

7.4.2.4 SENSITIVITY

Comparing the performance of Approach III.A and III.B when applied to dataset F_{pq_1} , given in Table 7.16, only marginal improvements are seen when making use of quantitative eigenvalue matchings. As a standalone technique, however, neither of the two approaches can be deemed adequately sensitive.

Table 7.16: Comparison of the performance metrics of dataset F_{pq_1} for Approach III.A and Approach III.B

Approach	r_{FP}	r_{FN}	r_{TP}	Accuracy	Isolability [%]	Isolation [%]
III.A	0.0	66.7	33.3	33.3	8.3	25.0
III.B	0.0	41.7	58.3	58.3	16.7	33.3

7.4.2.5 STORAGE AND COMPUTATIONAL REQUIREMENTS

Seeing as the differences between the qualitative and quantitative approaches are so small, the storage and computational requirements are regarded as being equal. Both approaches have slightly larger storage and computational requirements when compared to Approach II.B; however, as the methodology necessitates the calculation of the additional cost matrices for every operational condition considered (C_{oo}).

7.4.2.6 SUMMARISING REMARKS

By using quantitative instead of qualitative matchings, a significant improvement of the FDI performance was seen. Nevertheless, neither of the two approaches utilising more matching parameters could address the sensitivity issues seen in Approach II.B. A summary of the performance metrics is given in Table 7.17, and the cumulative performance is visually depicted in Figure 7.3.

Table 7.17: A summary of the performance metrics for Approach III.B

Fault set	r_{FP}	r_{FN}	r_{TP}	Accuracy	Isolability [%]	Isolation [%]
F_{pq_R}	0.0	8.3	91.7	92.3	84.6	84.6
F_{pq_1}	0.0	41.7	58.3	58.3	16.7	16.7

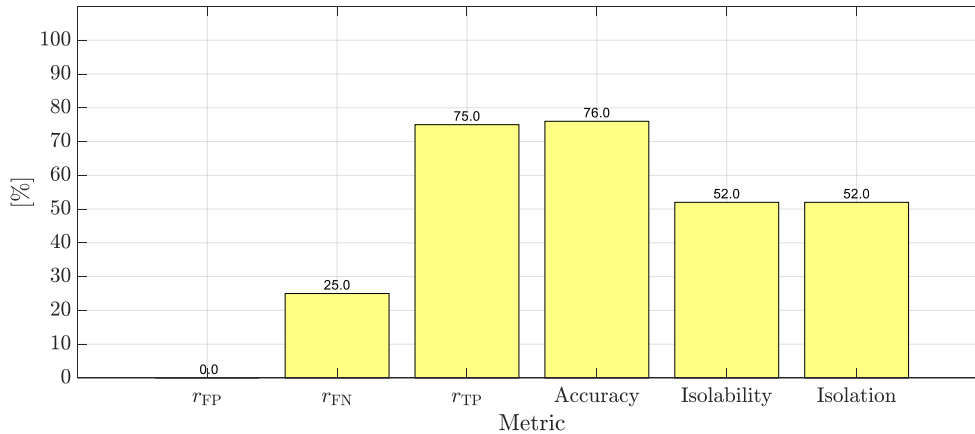


Figure 7.3: Visual depiction of the overall performance metrics of Approach III.B

The subsequent section details and compares the performance of all the approaches investigated in this study.

7.5 APPROACHES COMPARISON

To determine the best proposed approach, the discussed performance metrics are compared and deliberated. For ease of comparison, the various metrics throughout Chapter 5, Chapter 6, and Chapter 7 are summarised

in Table 7.18. Additionally, visual representations are given in Figure 7.4. Starting off with the detection accuracy, Approach I.A and II.B could both correctly detect all of the operational conditions within dataset F_{pq_R} . Although Approach III.B performed better than III.A, the techniques detected faults significantly worse than the most simplistic approach (I.A) investigated. When assessing dataset F_{pq_1} , the threshold approach (I.A) performed the best, being reasonably sensitive to the small 3 % faults; with only one fault undetected. The graph-based approaches (II.B, III.A, and III.B) all had sensitivity issues, as many faults were incorrectly classified as normal (NOC). Subsequently, the ranking of the approaches' cumulative detection capabilities - ranked best to worst - are I.A, II.B, III.B, and III.A.

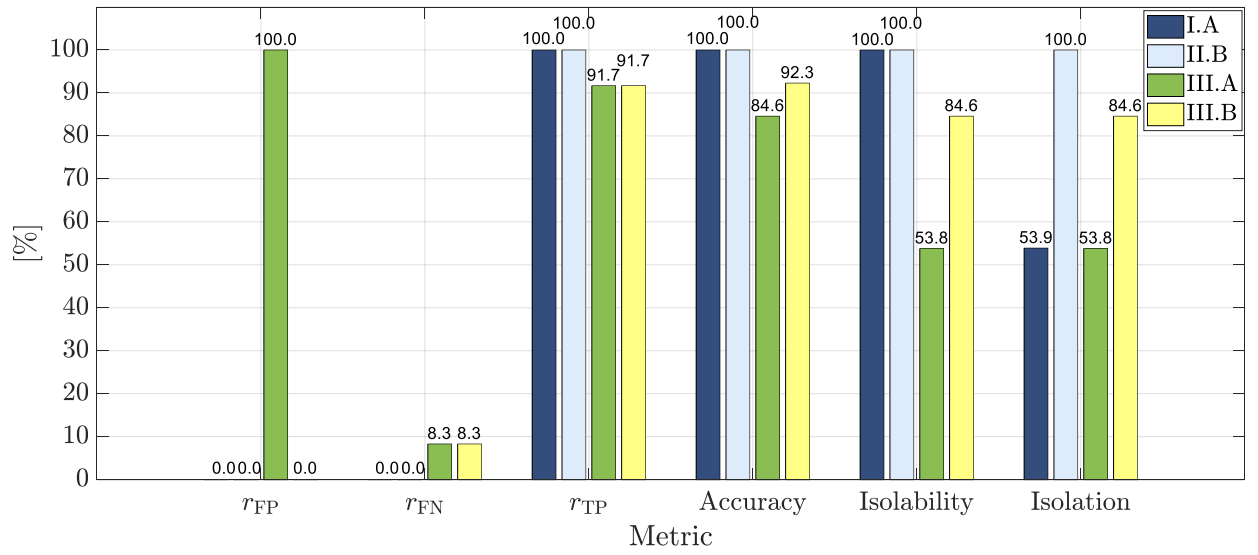
Considering the isolability of the approaches, both Approach I.A and II.B accomplished 100 % isolability for dataset F_{pq_R} . Once again, approaches III.A and III.B show poor performance, as meagre isolability percentages are attained. When assessing the isolability of dataset F_{pq_1} , the graph-based approaches all show very poor performance (incorrect NOC matchings). The ranking for the cumulative isolability is, therefore, I.A, II.B, III.B, and III.A.

Evaluating the isolation ability of the techniques, Approach II.B performed exceptionally well for dataset F_{pq_R} . All matchings correctly indicated the location of the fault. Not one of the other approaches could achieve this for F_{pq_R} . Once again, because Approach II.B categorised many operational conditions within F_{pq_1} as NOC, the performance saw a steep decrease. Notably, the only time Approach III.B out-performed II.B, was with the successful isolability and isolation of one additional operational condition (F_{11_1}). The cumulative isolation rankings are thus, II.B, I.A, III.B, and III.A.

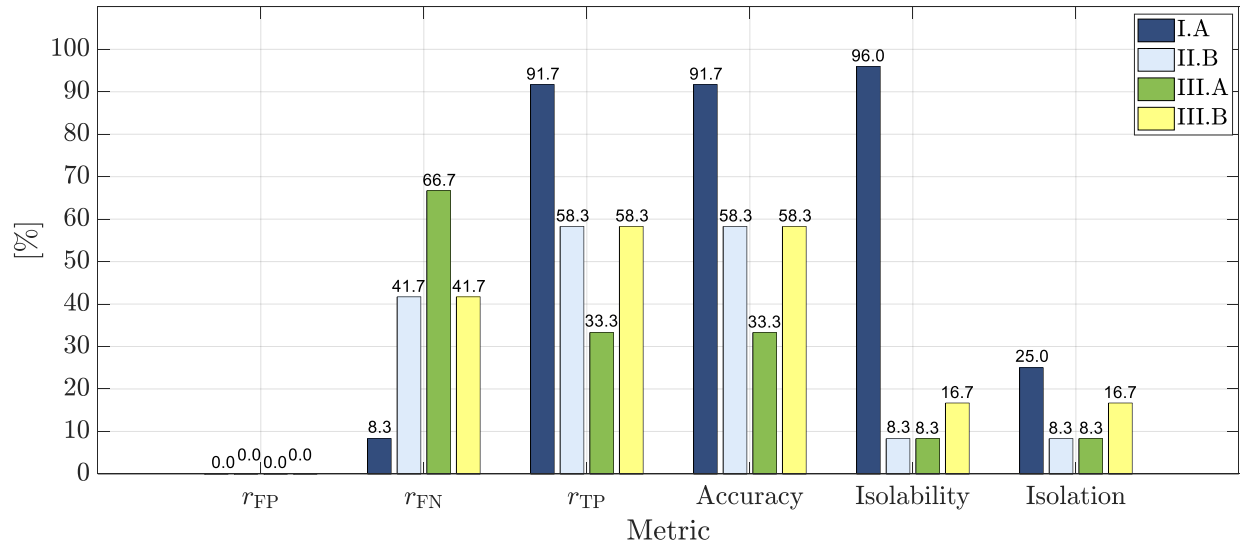
For interest's sake, the approaches' detected, isolable and isolated faults are visually summarised in Table 7.19 and Table 7.20.

Table 7.18: A summary of the performance metrics of the various approaches investigated

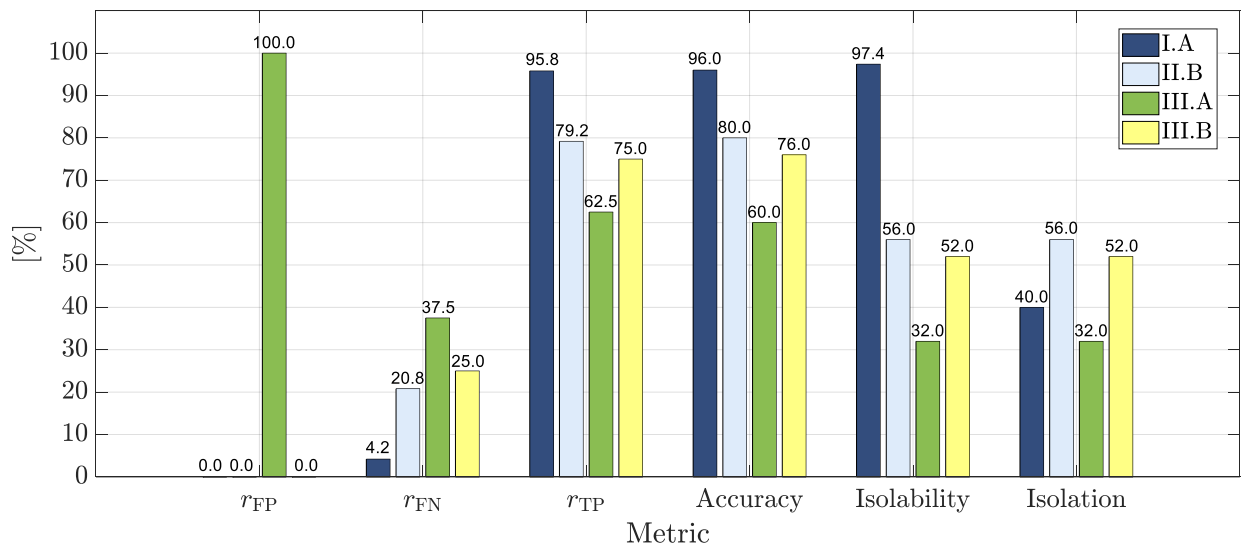
Fault set	Approach	Detection rates [%]				Isolability [%]	Isolation [%]
		r_{FP}	r_{FN}	r_{TP}	Accuracy		
F_{pq_R}	I.A	0.0	0.0	100.0	100.0	100.0	53.9
	II.B	0.0	0.0	100.0	100.0	100.0	100.0
	III.A	100.0	8.3	91.7	84.6	53.8	53.8
	III.B	0.0	8.3	91.7	92.3	84.6	84.6
F_{pq_1}	I.A	0.0	8.3	91.7	91.7	96.0	25.0
	II.B	0.0	41.7	58.3	58.3	8.3	8.3
	III.A	0.0	66.7	33.3	33.3	8.3	8.3
	III.B	0.0	41.7	58.3	58.3	16.7	16.7
Cumulative	I.A	0.0	4.2	95.8	96.0	97.4	40.0
	II.B	0.0	20.8	79.2	80.0	56.0	56.0
	III.A	100.0	37.5	62.5	60.0	32.0	32.0
	III.B	0.0	25.0	75.0	76.0	52.0	52.0



(a) F_{pq_R}



(b) F_{pq_1}



(c) Cumulative

Figure 7.4: Graphical representation of the approaches' performance for (a) F_{pq_R} , (b) F_{pq_1} and (c) cumulative

Table 7.19: Visual summary of the detection, isolability, and isolation of F_{pq_R} of all approaches

Fault ID		Detected	Isolable	Isolation		Detected	Isolable	Isolation		Detected	Isolable	Isolation		Detected	Isolable	Isolation
Normal ₁	Approach I.A	✓	✓	✓	Approach II.B	✓	✓	✓	Approach III.A	✗	✗	✗	Approach III.B	✓	✓	✓
F_{11_6}		✓	✓	✓		✓	✓	✓		✓	✗	✗		✓	✓	✓
F_{12_3}		✓	✓	✓		✓	✓	✓		✓	✗	✗		✓	✓	✓
F_{13_7}		✓	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓	✓
F_{14_2}		✓	✓	✗		✓	✓	✓		✓	✗	✗		✓	✓	✓
F_{21_4}		✓	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓	✓
F_{22_5}		✓	✓	✗		✓	✓	✓		✓	✓	✓		✓	✓	✓
F_{23_6}		✓	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓	✓
F_{24_2}		✓	✓	✓		✓	✓	✓		✓	✓	✓		✓	✓	✓
F_{31_7}		✓	✓	✗		✓	✓	✓		✓	✓	✓		✗	✗	✗
F_{32_2}		✓	✓	✗		✓	✓	✓		✗	✗	✗		✓	✓	✓
F_{33_5}		✓	✓	✗		✓	✓	✓		✓	✗	✗		✓	✓	✓
F_{34_3}		✓	✗	✗		✓	✓	✓		✓	✓	✓		✓	✗	✗

Table 7.20: Visual summary of the detection, isolability, and isolation of F_{pq_1} of all approaches

Fault ID		Detected	Isolable	Isolation		Detected	Isolable	Isolation		Detected	Isolable	Isolation		Detected	Isolable	Isolation
F_{11_1}	Approach I.A	✓	✓	✓	Approach II.B	✓	✗	✗	Approach III.A	✓	✓	✓	Approach III.B	✓	✓	✓
F_{12_1}		✓	✓	✓		✓	✗	✗		✗	✗	✗		✓	✗	✗
F_{13_1}		✓	✓	✗		✓	✓	✓		✓	✗	✗		✓	✓	✓
F_{14_1}		✓	✓	✗		✗	✗	✗		✓	✗	✗		✗	✗	✗
F_{21_1}		✓	✓	✗		✗	✗	✗		✗	✗	✗		✗	✗	✗
F_{22_1}		✓	✓	✗		✓	✗	✗		✗	✗	✗		✓	✗	✗
F_{23_1}		✗	✗	✗		✗	✗	✗		✗	✗	✗		✗	✗	✗
F_{24_1}		✓	✓	✓		✗	✗	✗		✗	✗	✗		✗	✗	✗
F_{31_1}		✓	✓	✗		✗	✗	✗		✗	✗	✗		✗	✗	✗
F_{32_1}		✓	✓	✗		✓	✗	✗		✗	✗	✗		✓	✗	✗
F_{33_1}		✓	✓	✗		✓	✗	✗		✓	✗	✗		✓	✗	✗
F_{34_1}		✓	✓	✗		✓	✗	✗		✗	✗	✗		✓	✗	✗

7.6 CONCLUSION

After contemplating the performance of the investigated approaches, the best performing approach seems to be the fixed-threshold (Approach I.A) when regarding the cumulative detection and isolability. The technique, however, has poor isolation capabilities; possibly because of the recycling stream propagating the fault throughout the system and the lack of an intelligent auxiliary stage(s) that could counter/account for the occurrence. The technique is, therefore, deemed quite suitable for detecting faults - should the threshold be adequately sensitive - but insufficiently capable of describing exact fault locations. As such, the proposed graph-matching approach (Approach II.B) would be better suited as a fault detection and isolation solution. The approach could successfully detect faults and pinpoint the exact location of the faults within the F_{pq_R} dataset. The approach, however, would need to be modified in order to achieve better sensitivity (indicated by the poor performance for dataset F_{pq_1}). Sadly, the utilisation of eigenvalues to effect database matchings produced unsatisfactory results. A quantitative strategy seemed to improve the eigendecomposition approach's performance only marginally compared to that of the qualitative matchings. The most probable reason for the better results would be that III.B only allows for the smallest quantitative differences in eigenvalues to be summed, compared to III.A where any eigenvalue falling beyond the 3σ boundary is assigned and summed.

CHAPTER 8

CONCLUSION

8.1 INTRODUCTION

This chapter serves as the conclusion to the thesis. Special attention is given to stipulate the contributions of the study. Additionally, some future work and relevant recommendations are discussed. The main focus of the research was to explore and compare the applicability (after some modifications) and performance of a few existing energy-based FDI approaches which endeavour to hybridise energy properties and structural information. The petrochemical process that was chosen - not yet seen within the FDD field - was a gas-to-liquids (GTL) process which includes one recycle stream. Table 8.1 recapitulates the approaches that were investigated. The subsequent sections will highlight the notable findings of the study.

Table 8.1: Summary of investigated energy-based FDI approaches and their details

Approach notation	Details	Classification	Based on work done by
Approach I.A	Fixed-threshold	Qualitative	[16]
Approach II.B	Graph-matching using D_C -value	Quantitative	-
Approach III.A	Graph-matching using eigenvalues	Qualitative	[18, 19]
Approach III.B	Graph-matching using eigenvalues	Quantitative	[17]

8.2 OUTCOME OF RESEARCH OBJECTIVES

8.2.1 GTL MODEL

A steady-state simulation model of a GTL process of representative complexity and scale was successfully created within Aspen HYSYS[®]. The GTL model could, with relative ease, be modified to simulate normal and pre-defined faulty conditions.

8.2.2 ENERGY CHARACTERISATION

To automatically calculate the desired physical and chemical exergy, *user variables* within HYSYS[®] were employed. After recording the computed exergy and energy flow data, the normal and faulty behaviour of the process could be characterised in terms of these energy properties.

8.2.3 FIXED-THRESHOLD APPROACH - APPROACH I.A

Utilising the compiled exergy data the first approach, the exergy-based, fixed-threshold approach detailed in the work of Marais [16], was applied. The applicability of the approach, particularly pertaining to a process that includes a recycle stream, was of primary interest. As suspected, the recycle stream proved to have a significant impact, as the faulty effects were propagated throughout the entire process. Nevertheless, the threshold approach's detection capabilities were found to be adequate, missing only one 3 % fault condition. The isolation performance which indicates the fault location, however, showed unfavourable results. This motivated the investigation of graph-based FDI approaches to determine whether the performance might improve, specifically regarding the isolation facilities.

8.2.4 GRAPH-MATCHING D_C -VALUE - APPROACH II.B

By using graph theory principles, one encapsulates not only structural information but also causal relations thereof. Moreover, the field allows for a wide range of mathematical methods in which to manipulate and eventually detect and isolate fault conditions. The two main graph-based approaches examined were one utilising a distance parameter (Approach II.B) and the other making use of eigendecomposition as proposed in [17–19] (Approach III.A and III.B). Notably, the utilisation of the distance parameter as a means to FDI was a novel proposition. Additionally, these approaches were based on the premise that a database with stored fault graphs was available. This is commonly seen in data-driven methods (historical data). Graph matching, a subdivision found within graph theory - which quantifies how dissimilar two compared graphs are - was used to match an operational condition's graph to the most similar fault graph stored within the database.

When regarding graph matching literature, matchings of graphs are achieved by firstly calculating a cost matrix and from the cost matrix a corresponding distance value (D_C). Note that the cost matrix is reduced to a single match-indicating parameter. The smaller the D_C -value, the more similar the two compared graphs are. The first graph-based approach, therefore, examines all operational conditions as matched to which corresponding database entries indicated by the smallest D_C -value. The proposed approach performed flawlessly when ≥ 8 % magnitude faults were assessed. Unfortunately, many 3 % faulty conditions were classified as being normal, signifying an issue with sensitivity. Nonetheless, a definitive improvement was seen in the isolation capability as compared to the threshold approach. In a bid to address the sensitivity, the eigendecomposition avenue was evaluated.

8.2.5 GRAPH-MATCHING EIGENVALUES - APPROACH III.A/B

As the eigendecomposition, specifically looking at the eigenvalues, provide a larger number of matching parameters, not just a single D_C -value, an improvement in performance was envisioned. In the works of [17–19] two different eigendecomposition routes are seen. Neser [18] and Uren et al. [19] express the eigenvalues in terms of qualitative signatures, whereas Van Graan [17] quantitatively visualised eigenvalues and their changes to achieve FDI. As such, for this study, the graph matching utilising eigenvalues were done

qualitatively (Approach III.A) and quantitatively (Approach III.B). Neither approaches seemed to improve the matchings. Moreover, the qualitative approach (Approach III.A) performed the worst of all the assessed approaches. It can therefore be argued that the best matching mechanism remains to be the distance parameter (D_C -value) used within the graph theory field. This is not to say that the eigenvalues, which characterises and describes a considered matrix, do not hold exploitable information. The eigenvalues should, however, not be used as matching mechanisms but rather be interpreted from a different perspective.

8.2.6 COMPARISON OF APPROACHES

When comparing the four approaches that were evaluated, the following aspects stood out. Even though Approach I.A had good detection and isolability performance, some isolation capabilities were lost due to the recycle stream's propagating nature. The energy-based, graph-based approach which utilises the distance parameter, Approach II.B, show excellent isolation when presented with $\geq 8\%$ fault magnitudes. Unfortunately, the approach performed inadequately when small magnitude faults are introduced (sensitivity issues). The utilisation of eigenvalues (Approach III.A/B) did not improve the matching capabilities, i.e., the FDI facilities. Consequently, one might need to utilise the information afforded by the eigenvalues in a different manner. Nevertheless, Approach II.B seem to show the most possibilities if the sensitivity issues could be addressed.

8.3 CONTRIBUTION

The following lists the contributions of the study:

- The development and comprehensive documentation of a steady-state simulation model of a gas-to-liquids (GTL) process plant that could lend itself to being used as an alternative larger-scale benchmark process for the purpose of testing and comparing proposed FDD techniques. Additionally, the *user variables* that automatically calculate the exergy within HYSYS[®] were expanded to account for various stream phases.
- The analysis of the applicability of an exergy-based, fixed-threshold approach (detailed in [16]) when applied to a petrochemical process that includes a recycle stream.
- The evaluation of the feasibility and performance of energy-based approaches that make use of energy properties and graph-based fundamentals. The two avenues of graph matching investigated, being:
 - A graph matching approach which utilises a single distance parameter.
 - A graph matching approach which employs eigenvalues (eigendecomposition).

8.4 FUTURE WORK AND RECOMMENDATIONS

8.4.1 GTL AS BENCHMARK PROCESS

With every newly developed FDD technique, the base system and case studies might vary, complicating the direct comparison of various FDD approaches. The issue can be addressed by making use of a known benchmark process. One of the most popular large-scale benchmark processes seen within the FDD field is the Tennessee Eastman Process (TEP). Because of its proprietary nature, much of the process and exact details are obfuscated. Only a few different model platforms exist; of which only one model is found in the form of a commercial process simulation. Process insight and data are, therefore, very limited. Although many researchers make use of the TEP to demonstrate their proposed FDI techniques, a more accessible benchmark process with representative complexity would immensely benefit the FDD community. Even though some aspects of the gas-to-liquids (GTL) process are also proprietary, a large number of literary work sufficiently documents the required particulars. Since the process functions, operating points, and viable fault conditions are comprehensively detailed in Chapter 3 and Appendix A; it is recommended that the GTL model developed for this study be used as an alternative benchmark process (within the petrochemical domain). Furthermore, as the simulation was done in Aspen HYSYS[®], minor modifications would be necessary to convert the steady-state model to a dynamic one.

8.4.2 APPROACH SENSITIVITY

Seeing as the FDI approach, utilising the D_C -value, performed so well for larger magnitude faults, it would be beneficial to address some of the sensitivity issues exhibited. The author would propose exploring the effects of combining the threshold and graph-based approaches. As an example, will the matchings improve if (1) the threshold approach, detecting a faulty condition, is used to exclude the normal condition (NOC) from the database and then (2) the graph matching approach is used to determine a matching fault condition?

8.4.3 MULTIPLE FAULTS

This study only evaluated single faults. A common practice within the FDD field is to assess a proposed technique's ability to handle multiple faults, and as such, should be included in future explorations.

8.4.4 DYNAMIC SYSTEM

The GTL process, which is presumed to be relatively complex scale-wise, was observed at steady-state conditions only. As the employment of energy-based FDI is relatively new, the concepts and limitations thereof were firstly evaluated and compared. Therefore, the steady-state conditions were deemed sufficient for testing purposes. It should, however, be emphasised that any real process plant (and its behaviours) would vastly differ from a modelled one; and a steady-state model at that. Variability that impacts a process plant comes in many forms, including but not limited to uncertainty, process drift, environmental factors, and feedstock variations. Fundamentally, then, the investigated approaches might need to (1) be modified to account for these influences and (2) be applied to a dynamic version of the process featuring some of

these real plant attributes. Additionally, with the inclusion of dynamic behaviour, the 3 % faults might not be distinguishable from dynamic disturbances and would probably generate false positives. Thus, the magnitude of the faults would need to be re-examined.

8.4.5 INCLUSION OF SENSOR NOISE

For this study, some simulation solver variations (numerical noise) were observed and even utilised in some instances. For the most part, the variations were found to be small enough to be negligible. In actual systems, however, sensor noise has larger and more distinct influences. The effect of sensor noise should, therefore, be included and evaluated in future work. Seeing as the fixed-threshold (Approach I.A) is based on smaller numerical noise, it is expected that the inclusion of sensor noise would drastically effect the approach's ability to detect fault-free conditions. Similar to the inclusion of dynamic disturbances, this will result in an increase of false positives. Conversely, the additional variation might improve the graph-based approaches' performance when considering small magnitude faults.

8.5 CLOSURE

The thesis endeavoured to extend, evaluate, and compare some of the existing energy-based FDI approaches as applied to a petrochemical process, specifically a GTL process. The value of using energy as a unifying domain parameter is verified once more. The graph-based approaches have the added benefit of retaining not only energy and structural data - such as the threshold approach - but also to encapsulate causal relations. Of the graph-based approaches assessed, the graph-matching approach utilising the D_C -value seems to have the most potential to be explored and modified, in order to improve performance when confronted with smaller magnitude faults. Although some promising insight was gained regarding energy-based FDI approaches, many aspects still need to be explored and addressed.

BIBLIOGRAPHY

- [1] T. Kletz, *What Went Wrong? Case Histories of Process Plant Disasters*, 4th ed. Elsevier Science, 1998.
- [2] V. Venkatasubramanian, R. Raghunathan, K. Yin, and S. Kavuri, “A review of process fault detection and diagnosis: Part I: Quantitative model-based methods,” *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 293–311, 2003. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(02\)00160-6](https://doi.org/10.1016/S0098-1354(02)00160-6)
- [3] Z. Gao, C. Cecati, and S. X. Ding, “A survey of fault diagnosis and fault-tolerant techniques—part II: Fault diagnosis with knowledge-based and hybrid/active approaches,” *IEEE Transactions on Industrial Electronics*, vol. 62, p. 3768–3774, 2015.
- [4] S. J. Qin, “Data-driven fault detection and diagnosis for complex industrial processes,” *IFAC Proceedings Volumes*, vol. 42, no. 8, pp. 1115–1125, 2009. [Online]. Available: <https://doi.org/10.3182/20090630-4-ES-2003.00184>
- [5] M. Z. Sheriff, C. Botre, M. Mansouri, H. Nounou, M. Nounou, and M. N. Karim, “Process monitoring using data-based fault detection techniques: Comparative studies,” in *Fault Diagnosis and Detection*. InTech: Münchenstein, Switzerland, 2017.
- [6] F. Harrou, M. Madakyaru, and Y. Sun, “Improved nonlinear fault detection strategy based on the Hellinger distance metric: Plug flow reactor monitoring,” *Energy and Buildings*, vol. 143, pp. 149–161, 2017. [Online]. Available: <https://doi.org/10.1016/j.enbuild.2017.03.033>
- [7] K. Villez and J. Habermacher, “Shape anomaly detection for process monitoring of a sequencing batch reactor,” *Computers & Chemical Engineering*, vol. 91, pp. 365–379, 2016. [Online]. Available: <https://doi.org/10.1016/j.compchemeng.2016.04.012>
- [8] K. Tidriri, N. Chatti, S. Verron, and T. Tiplica, “Model-based fault detection and diagnosis of complex chemical processes: A case study of the Tennessee Eastman process,” *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, vol. 232, no. 6, pp. 742–760, 2018. [Online]. Available: <https://doi.org/10.1177/0959651818764510>
- [9] Z. Gao, C. Cecati, and S. X. Ding, “A survey of fault diagnosis and fault-tolerant techniques—part I: Fault diagnosis with model-based and signal-based approaches,” *IEEE Transactions on Industrial Electronics*, vol. 62, no. 6, pp. 3757–3767, 2015. [Online]. Available: 10.1109/TIE.2015.2417501

- [10] C. P. Du Rand, G. Van Schoor, and C. Nieuwoudt, "Enthalpy-entropy graph approach for the classification of faults in the main power system of a closed Brayton cycle HTGR," *Annals of Nuclear Energy*, vol. 36, no. 6, pp. 703–711, 2009. [Online]. Available: <https://doi.org/10.1016/j.anucene.2009.02.012>
- [11] H. Marais, G. van Schoor, and K. R. Uren, "An Energy-based approach to condition monitoring of industrial processes," IFAC. Elsevier, 2015, pp. 772–777. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2015.09.620>
- [12] —, "The merits of exergy-based fault detection in petrochemical processes," *Journal of Process Control*, vol. 74, pp. 110–119, 2019. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2017.11.005>
- [13] K. Tidriri, N. Chatti, S. Verron, and T. Tiplica, "Bridging data-driven and model-based approaches for process fault diagnosis and health monitoring: A review of researches and future challenges," *Annual Reviews in Control*, vol. 42, pp. 63–81, 2016. [Online]. Available: <https://doi.org/10.1016/j.arcontrol.2016.09.008>
- [14] M. T. Amin, S. Imtiaz, and F. Khan, "Process system fault detection and diagnosis using a hybrid technique," *Chemical Engineering Science*, vol. 189, pp. 191–211, 2018. [Online]. Available: <https://doi.org/10.1016/j.ces.2018.05.045>
- [15] R. S. Martins, M. R. Vale, and A. L. Maitelli, "Hybrid methods for detection and identification of faults in dynamic systems," *Asian Journal of Control*, vol. 17, no. 5, pp. 1831–1847, 2015. [Online]. Available: <https://doi.org/10.1002/asjc.1039>
- [16] H. Marais, "An energy-based approach to condition monitoring of an auto-thermal reformer," Ph.D. dissertation, Computer and Electronic Engineering, North-West University, 2015.
- [17] S. Van Graan, G. Van Schoor, and K. Uren, "Graph matching as a means to energy-visualisation of a counter-flow heat exchanger for the purpose of fault diagnosis," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 2842–2847, 2017. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2017.08.637>
- [18] H. Naser, "Energy based visualisation of a Brayton cycle power conversion unit for the purpose of condition monitoring," Ph.D. dissertation, Computer and Electronic Engineering, North-West University, 2019.
- [19] K. R. Uren, G. van Schoor, and L. Aurret, "An energy-attributed graph approach for the purposes of FDI in a heated two-tank system," *IFAC-PapersOnLine*, vol. 52, no. 14, pp. 159–164, 2019. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2019.09.181>
- [20] X. Hao, M. E. Djatmiko, Y. Xu, Y. Wang, J. Chang, and Y. Li, "Simulation analysis of a GTL process using ASPEN plus," *Chemical Engineering & Technology*, vol. 31, no. 2, pp. 188–196, 2008. [Online]. Available: <https://doi.org/10.1002/ceat.200700336>

- [21] A. Rafiee and M. Hillestad, "Optimal design and operation of a gas-to-liquid process," *CHEMICAL ENGINEERING TRANSACTIONS*, vol. 21, pp. 1393–1398, 2010.
- [22] B. Bao, M. M. El-Halwagi, and N. O. Elbashir, "Simulation, integration, and economic analysis of gas-to-liquid processes," *Fuel Processing Technology*, vol. 91, no. 7, pp. 703–713, 2010. [Online]. Available: <https://doi.org/10.1016/j.fuproc.2010.02.001>
- [23] M. Panahi, A. Rafiee, S. Skogestad, and M. Hillestad, "A natural gas to liquids process model for optimal operation," *Industrial & Engineering Chemistry Research*, vol. 51, no. 1, pp. 425–433, 2011. [Online]. Available: <https://doi.org/10.1021/ie2014058>
- [24] K. T. Knutsen, "Modelling and optimization of a Gas-to-Liquid plant," Master's thesis, Norwegian University of Science and Technology, 2013.
- [25] I. Dincer and M. A. Rosen, *Exergy: energy, environment and sustainable development*, 2nd ed. Elsevier, 2013.
- [26] K. Severson, P. Chaiwatanodom, and R. D. Braatz, "Perspectives on process monitoring of industrial systems," *Annual Reviews in Control*, vol. 42, pp. 190–200, 2016. [Online]. Available: <https://doi.org/10.1016/j.arcontrol.2016.09.001>
- [27] G. Stanley, "A guide to fault detection and diagnosis," accessed 06-09-2018. [Online]. Available: <https://gregstanleyandassociates.com/whitepapers/FaultDiagnosis/faultdiagnosis.htm>
- [28] J. J. Gertler, *Fault Detection and Diagnosis in Engineering Systems*. Marcel Dekker, Inc., 1998.
- [29] R. Isermann, "Process fault detection based on modeling and estimation methods—A survey," *Automatica*, vol. 20, no. 4, pp. 387–404, 1984. [Online]. Available: [https://doi.org/10.1016/0005-1098\(84\)90098-0](https://doi.org/10.1016/0005-1098(84)90098-0)
- [30] S. X. Ding, *Model-based fault diagnosis techniques: Design schemes, algorithms, and tools*. Springer Science & Business Media, 2008.
- [31] V. Venkatasubramanian, R. Raghunathan, K. Yin, and S. Kavuri, "A review of process fault detection and diagnosis: Part II: Qualitative models and search strategies," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(02\)00161-8](https://doi.org/10.1016/S0098-1354(02)00161-8)
- [32] —, "A review of process fault detection and diagnosis: Part III: Process history based methods," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 327–346, 2003. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(02\)00162-X](https://doi.org/10.1016/S0098-1354(02)00162-X)
- [33] B. Ould-Bouamama, G. Biswas, R. Loureiro, and R. Merzouki, "Graphical methods for diagnosis of dynamic systems: Review," *Annual Reviews in Control*, vol. 38, no. 2, pp. 199–219, 2014. [Online]. Available: <https://doi.org/10.1016/j.arcontrol.2014.09.004>

- [34] F. Pierri, G. Paviglianiti, F. Caccavale, and M. Mattei, "Observer-based sensor fault detection and isolation for chemical batch reactors," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 8, pp. 1204–1216, 2008. [Online]. Available: <https://doi.org/10.1016/j.engappai.2008.02.002>
- [35] J. Zarei and E. Shokri, "Robust sensor fault detection based on nonlinear unknown input observer," *Measurement*, vol. 48, pp. 355–367, 2014. [Online]. Available: <https://doi.org/10.1016/j.measurement.2013.11.015>
- [36] L. Rusinov, N. Vorobiev, and V. Kurkina, "Fault diagnosis in chemical processes and equipment with feedbacks," *Chemometrics and Intelligent Laboratory Systems*, vol. 126, pp. 123–128, 2013. [Online]. Available: <https://doi.org/10.1016/j.chemolab.2013.03.015>
- [37] E. M. Cimpoesu, B. D. Ciubotaru, and D. Stefanoiu, "Fault Detection and Diagnosis Using Parameter Estimation with Recursive Least Squares," in *Control Systems and Computer Science (CSCS), 19th International Conference*. IEEE, 2013, pp. 18–23.
- [38] T. Escobet and L. Travé-Massuyès, "Parameter estimation methods for fault detection and isolation," in *Proceedings of the 12th International Workshop on Principles of Diagnosis (The BRIDGE Workshop)*, vol. 40. Citeseer, 2001.
- [39] R. Isermann, "Model-based fault-detection and diagnosis - status and applications," *Annual Reviews in Control*, vol. 29, no. 1, pp. 71–85, 2005. [Online]. Available: <https://doi.org/10.1016/j.arcontrol.2004.12.002>
- [40] B. Ould-Bouamama, R. El Harabi, M. N. Abdelkrim, and M. B. Gayed, "Bond graphs for the diagnosis of chemical processes," *Computers & Chemical Engineering*, vol. 36, pp. 301–324, 2012. [Online]. Available: <https://doi.org/10.1016/j.compchemeng.2011.07.008>
- [41] S. Katipamula and M. R. Brambley, "Methods for fault detection, diagnostics, and prognostics for building systems—a review, part I," *Hvac&R Research*, vol. 11, no. 1, pp. 3–25, 2005.
- [42] M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian, "A systematic framework for the development and analysis of signed digraphs for chemical processes. 1. algorithms and analysis," *Industrial & Engineering Chemistry Research*, vol. 42, no. 20, pp. 4789–4810, 2003. [Online]. Available: <https://doi.org/10.1021/ie020644a>
- [43] N. H. Ulerich and G. J. Powers, "On-line hazard aversion and fault diagnosis in chemical processes: the digraph+fault-tree method," *IEEE Transactions on Reliability*, vol. 37, no. 2, pp. 171–177, 1988. [Online]. Available: <https://doi.org/10.1109/24.3738>
- [44] J. De Kleer and J. S. Brown, "A qualitative physics based on confluences," *Artificial Intelligence*, vol. 24, no. 1-3, pp. 7–83, 1984. [Online]. Available: [https://doi.org/10.1016/0004-3702\(84\)90037-7](https://doi.org/10.1016/0004-3702(84)90037-7)
- [45] P. Silva, A. Carvalho, H. Gabbar, P. Vieira, and C. Costa, "Fault diagnosis in transmission lines based on leakage current and qualitative trend analysis," in *2017 International Conference on Promising Electronic Technologies (ICPET)*. IEEE, 2017, pp. 87–92.

- [46] S. Ding, P. Zhang, B. Huang, and E. Ding, "Subspace method aided data-driven design of observer based fault detection systems," *IFAC Proceedings Volumes*, vol. 38, no. 1, pp. 167–172, 2005. [Online]. Available: <https://doi.org/10.3182/20050703-6-CZ-1902.01830>
- [47] L. H. Chiang and R. D. Braatz, "Process monitoring using causal map and multivariate statistics: fault detection and identification," *Chemometrics and Intelligent Laboratory Systems*, vol. 65, no. 2, pp. 159–178, 2003. [Online]. Available: [https://doi.org/10.1016/S0169-7439\(02\)00140-5](https://doi.org/10.1016/S0169-7439(02)00140-5)
- [48] G. Lee, S.-O. Song, and E. S. Yoon, "Multiple-fault diagnosis based on system decomposition and dynamic PLS," *Industrial & Engineering Chemistry Research*, vol. 42, no. 24, pp. 6145–6154, 2003. [Online]. Available: <https://doi.org/10.1021/ie030084v>
- [49] A. Khoukhi and M. H. Khalid, "Hybrid computing techniques for fault detection and isolation, a review," *Computers & Electrical Engineering*, vol. 43, pp. 17–32, 2015. [Online]. Available: <https://doi.org/10.1016/j.compeleceng.2014.12.015>
- [50] A. Berton and D. Hodouin, "Linear and bilinear fault detection and diagnosis based on mass and energy balance equations," *Control Engineering Practice*, vol. 11, no. 1, pp. 103–113, 2003. [Online]. Available: [https://doi.org/10.1016/S0967-0661\(02\)00116-8](https://doi.org/10.1016/S0967-0661(02)00116-8)
- [51] D. Theilliol, H. Noura, D. Sauter, and F. Hamelin, "Sensor fault diagnosis based on energy balance evaluation: Application to a metal processing," *ISA Transactions*, vol. 45, no. 4, pp. 603–610, 2006. [Online]. Available: [https://doi.org/10.1016/S0019-0578\(07\)60235-3](https://doi.org/10.1016/S0019-0578(07)60235-3)
- [52] W. Chen, "Fault detection and isolation in nonlinear systems: observer and energy-balance based approaches," Ph.D. dissertation, Universitätsbibliothek Duisburg-Essen, 2011.
- [53] K. R. Uren and G. van Schoor, "Energy-based visualisation of a counter-flow heat exchanger for the purpose of fault identification." Trondheim, Norway: Symposium on Dynamics and Control of Process Systems, including Biosystems, June 2016, pp. 19–24. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2016.07.210>
- [54] G. van Schoor, K. R. Uren, M. A. van Wyk, P. A. van Vuuren, and P. Carel, "An energy perspective on modelling, supervision, and control of large-scale industrial systems: Survey and framework," in *World Congress*, vol. 19, no. 1, 2014, pp. 6692–6703. [Online]. Available: <https://doi.org/10.3182/20140824-6-ZA-1003.02190>
- [55] S. A. Patel and A. K. Kamrani, "Intelligent decision support system for diagnosis and maintenance of automated systems," *Computers & Industrial Engineering*, vol. 30, no. 2, pp. 297–319, 1996. [Online]. Available: [https://doi.org/10.1016/0360-8352\(95\)00173-5](https://doi.org/10.1016/0360-8352(95)00173-5)
- [56] T. A. Reddy, "Formulation of a generic methodology for assessing fdd methods and its specific adoption to large chillers," *ASHRAE Transactions*, vol. 113, no. 2, 2007.

- [57] T. Kurtoglu, O. J. Mengshoel, and S. Poll, "A framework for systematic benchmarking of monitoring and diagnostic systems," in *2008 International Conference on Prognostics and Health Management*. IEEE, 2008, pp. 1–13.
- [58] L. Ming and J. Zhao, "Review on chemical process fault detection and diagnosis," in *2017 6th International Symposium on Advanced Control of Industrial Processes (AdCONIP)*. IEEE, 2017, pp. 457–462.
- [59] Z. Ge, "Review on data-driven modeling and monitoring for plant-wide industrial processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 171, pp. 16–25, 2017. [Online]. Available: <https://doi.org/10.1016/j.chemolab.2017.09.021>
- [60] N. Scenna, "Some aspects of fault diagnosis in batch processes," *Reliability Engineering & System Safety*, vol. 70, no. 1, pp. 95–110, 2000. [Online]. Available: [https://doi.org/10.1016/S0951-8320\(00\)00049-1](https://doi.org/10.1016/S0951-8320(00)00049-1)
- [61] P. Nomikos and J. F. MacGregor, "Monitoring batch processes using multiway principal component analysis," *AIChE Journal*, vol. 40, no. 8, pp. 1361–1375, 1994. [Online]. Available: <https://doi.org/10.1002/aic.690400809>
- [62] D. S. Lee, J. M. Park, and P. A. Vanrolleghem, "Adaptive multiscale principal component analysis for on-line monitoring of a sequencing batch reactor," *Journal of Biotechnology*, vol. 116, no. 2, pp. 195–210, 2005. [Online]. Available: <https://doi.org/10.1016/j.jbiotec.2004.10.012>
- [63] T. Kourti, P. Nomikos, and J. F. MacGregor, "Analysis, monitoring and fault diagnosis of batch processes using multiblock and multiway PLS," *Journal of Process Control*, vol. 5, no. 4, pp. 277–284, 1995. [Online]. Available: [https://doi.org/10.1016/0959-1524\(95\)00019-M](https://doi.org/10.1016/0959-1524(95)00019-M)
- [64] A. Benkouider, J. Buvat, J. Cosmao, and A. Saboni, "Fault detection in semi-batch reactor using the EKF and statistical method," *Journal of Loss Prevention in the Process Industries*, vol. 22, no. 2, pp. 153–161, 2009. [Online]. Available: <https://doi.org/10.1016/j.jlp.2008.11.006>
- [65] A. Benkouider, R. Kessas, A. Yahiaoui, J. Buvat, and S. Guella, "A hybrid approach to faults detection and diagnosis in batch and semi-batch reactors by using EKF and neural network classifier," *Journal of Loss Prevention in the Process Industries*, vol. 25, no. 4, pp. 694–702, 2012. [Online]. Available: <https://doi.org/10.1016/j.jlp.2012.03.005>
- [66] P. Nomikos and J. F. MacGregor, "Multivariate SPC charts for monitoring batch processes," *Technometrics*, vol. 37, no. 1, pp. 41–59, 1995.
- [67] F. Caccavale, F. Pierri, M. Iamarino, and V. Tufano, "An integrated approach to fault diagnosis for a class of chemical batch processes," *Journal of Process Control*, vol. 19, no. 5, pp. 827–841, 2009. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2008.11.003>

- [68] T. Umeda, T. Kuriyama, E. O'shima, and H. Matsuyama, "A graphical approach to cause and effect analysis of chemical processing systems," *Chemical Engineering Science*, vol. 35, no. 12, pp. 2379–2388, 1980. [Online]. Available: [https://doi.org/10.1016/0009-2509\(80\)85051-2](https://doi.org/10.1016/0009-2509(80)85051-2)
- [69] J. Shiozaki, H. Matsuyama, E. O'shima, and M. Iri, "An improved algorithm for diagnosis of system failures in the chemical process," *Computers & Chemical Engineering*, vol. 9, no. 3, pp. 285–293, 1985. [Online]. Available: [https://doi.org/10.1016/0098-1354\(85\)80006-5](https://doi.org/10.1016/0098-1354(85)80006-5)
- [70] M. Iri, K. Aoki, E. O'Shima, and H. Matsuyama, "An algorithm for diagnosis of system failures in the chemical process," *Computers & Chemical Engineering*, vol. 3, pp. 489–493, 1979.
- [71] S. Grantham and L. Ungar, "A first principles approach to automated troubleshooting of chemical plants," *Computers & Chemical Engineering*, vol. 14, no. 7, pp. 783–798, 1990. [Online]. Available: [https://doi.org/10.1016/0098-1354\(90\)87086-5](https://doi.org/10.1016/0098-1354(90)87086-5)
- [72] P. Prasad, J. Davis, Y. Jirapinyo, J. R. Josephson, and M. Bhalodia, "Structuring diagnostic knowledge for large-scale process systems," *Computers & Chemical Engineering*, vol. 22, no. 12, pp. 1897–1905, 1998. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(98\)00227-0](https://doi.org/10.1016/S0098-1354(98)00227-0)
- [73] T. Ramesh, S. Shum, and J. Davis, "A structured framework for efficient problem solving in diagnostic expert systems," *Computers & Chemical Engineering*, vol. 12, pp. 891–902, 1988. [Online]. Available: [https://doi.org/10.1016/0098-1354\(88\)87016-9](https://doi.org/10.1016/0098-1354(88)87016-9)
- [74] B. Bakshi and G. Stephanopoulos, "Temporal representation of process trends for diagnosis and control," *IFAC Proceedings Volumes*, vol. 25, no. 4, pp. 109–114, 1992. [Online]. Available: [https://doi.org/10.1016/S1474-6670\(17\)50225-6](https://doi.org/10.1016/S1474-6670(17)50225-6)
- [75] J. Hoskins, K. Kaliyur, and D. M. Himmelblau, "Fault diagnosis in complex chemical plants using artificial neural networks," *AIChE Journal*, vol. 37, no. 1, pp. 137–141, 1991. [Online]. Available: <https://doi.org/10.1002/aic.690370112>
- [76] K. Watanabe and S. Hirota, "Incipient diagnosis of multiple faults in chemical processes via hierarchical artificial neural network," in *Proceedings IECON'91: 1991 International Conference on Industrial Electronics, Control and Instrumentation*. IEEE, 1991, pp. 1500–1505.
- [77] V. Venkatasubramanian, R. Vaidyanathan, and Y. Yamamoto, "Process fault detection and diagnosis using neural networks—I. Steady-state processes," *Computers & Chemical Engineering*, vol. 14, no. 7, pp. 699–712, 1990. [Online]. Available: [https://doi.org/10.1016/0098-1354\(90\)87081-Y](https://doi.org/10.1016/0098-1354(90)87081-Y)
- [78] S. Joe Qin, "Recursive PLS algorithms for adaptive data modeling," *Computers & Chemical Engineering*, vol. 22, no. 4, pp. 503–514, 1998. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(97\)00262-7](https://doi.org/10.1016/S0098-1354(97)00262-7)
- [79] S. Dash, S. Kantharao, R. Rengaswamy, and V. Venkatasubramanian, "Application and evaluation of linear/restricted nonlinear observers to a nonlinear CSTR," in *Computer Aided Chemical Engineering*, 2001, vol. 9, pp. 853–858. [Online]. Available: [https://doi.org/10.1016/S1570-7946\(01\)80136-X](https://doi.org/10.1016/S1570-7946(01)80136-X)

- [80] C. C. Chang and C. C. Yu, "On-line fault diagnosis using the signed directed graph," *Industrial & Engineering Chemistry Research*, vol. 29, no. 7, pp. 1290–1299, 1990. [Online]. Available: <https://doi.org/10.1021/ie00103a031>
- [81] R. Li and X. Wang, "Qualitative/quantitative simulation of process temporal behavior using clustered fuzzy digraphs," *AIChE Journal*, vol. 47, no. 4, pp. 906–919, 2001. [Online]. Available: <https://doi.org/10.1002/aic.690470413>
- [82] J. Zhang and P. Roberts, "Process fault diagnosis with diagnostic rules based on structural decomposition," *Journal of Process Control*, vol. 1, no. 5, pp. 259–269, 1991. [Online]. Available: [https://doi.org/10.1016/0959-1524\(91\)85017-D](https://doi.org/10.1016/0959-1524(91)85017-D)
- [83] J. Yu and M. M. Rashid, "A novel dynamic bayesian network-based networked process monitoring approach for fault detection, propagation identification, and root cause diagnosis," *AIChE Journal*, vol. 59, no. 7, pp. 2348–2365, 2013. [Online]. Available: <https://doi.org/10.1002/aic.14013>
- [84] Z. Zhang and F. Dong, "Fault detection and diagnosis for missing data systems with a three time-slice dynamic bayesian network approach," *Chemometrics and Intelligent Laboratory Systems*, vol. 138, pp. 30–40, 2014. [Online]. Available: <https://doi.org/10.1016/j.chemolab.2014.07.009>
- [85] S. W. Choi, C. Lee, J.-M. Lee, J. H. Park, and I.-B. Lee, "Fault detection and identification of nonlinear processes based on kernel PCA," *Chemometrics and Intelligent Laboratory Systems*, vol. 75, no. 1, pp. 55–67, 2005. [Online]. Available: <https://doi.org/10.1016/j.chemolab.2004.05.001>
- [86] K. Ghosh, S. Natarajan, and R. Srinivasan, "Hierarchically Distributed Fault Detection and Identification through Dempster-Shafer Evidence Fusion," *Industrial & Engineering Chemistry Research*, vol. 50, no. 15, pp. 9249–9269, 2011. [Online]. Available: <https://doi.org/10.1021/ie2003329>
- [87] H. Ji, X. He, J. Shang, and D. Zhou, "Incipient fault detection with smoothing techniques in statistical process monitoring," *Control Engineering Practice*, vol. 62, pp. 11–21, 2017. [Online]. Available: <https://doi.org/10.1016/j.conengprac.2017.03.001>
- [88] C. Botre, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Multiscale PLS-based GLRT for fault detection of chemical processes," *Journal of Loss Prevention in the Process Industries*, vol. 46, pp. 143–153, 2017. [Online]. Available: <https://doi.org/10.1016/j.jlp.2017.01.008>
- [89] N. Wilcox and D. Himmelblau, "The possible cause and effect graphs (PCEG) model for fault diagnosis-II. applications," *Computers & Chemical Engineering*, vol. 18, no. 2, pp. 117–127, 1994. [Online]. Available: [https://doi.org/10.1016/0098-1354\(94\)80132-0](https://doi.org/10.1016/0098-1354(94)80132-0)
- [90] T. Sorsa, H. N. Koivo, and H. Koivisto, "Neural networks in process fault diagnosis," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 4, pp. 815–825, 1991. [Online]. Available: <https://doi.org/10.1109/21.108299>

- [91] S. Narasimhan, P. Vachhani, and R. Rengaswamy, "New nonlinear residual feedback observer for fault diagnosis in nonlinear systems," *Automatica*, vol. 44, no. 9, pp. 2222–2229, 2008. [Online]. Available: <https://doi.org/10.1016/j.automatica.2007.12.020>
- [92] Y. H. S. D. G. Reklitis and V. Venkatasubramanian, "EKF based estimator for FDI in the model iv FCCU," in *IFAC conference SAFEPROCESS*, 2000.
- [93] H. Vedam and V. Venkatasubramanian, "Signed digraph based multiple fault diagnosis," *Computers & Chemical Engineering*, vol. 21, pp. S655–S660, 1997. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(97\)87577-1](https://doi.org/10.1016/S0098-1354(97)87577-1)
- [94] T. Ramesh, J. Davis, and G. Schwenzer, "Knowledge-based diagnostic systems for continuous process operations based upon the task framework," *Computers & Chemical Engineering*, vol. 16, no. 2, pp. 109–127, 1992.
- [95] H. Vedam and V. Venkatasubramanian, "A wavelet theory-based adaptive trend analysis system for process monitoring and diagnosis," in *Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041)*, vol. 1. IEEE, 1997, pp. 309–313.
- [96] X. Wang, B. Chen, S. Yang, and C. McGreavy, "Application of wavelets and neural networks to diagnostic system development, 2, an integrated framework and its application," *Computers & Chemical Engineering*, vol. 23, no. 7, pp. 945–954, 1999. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(99\)00260-4](https://doi.org/10.1016/S0098-1354(99)00260-4)
- [97] U. Kruger, X. Wang, Q. Chen, and S. Qin, "An alternative PLS algorithm for the monitoring of industrial process," in *Proceedings of the 2001 American Control Conference.(Cat. No. 01CH37148)*, vol. 6. IEEE, 2001, pp. 4455–4459.
- [98] Y. Qian, X. Li, Y. Jiang, and Y. Wen, "An expert system for real-time fault diagnosis of complex chemical processes," *Expert Systems with Applications*, vol. 24, no. 4, pp. 425–432, 2003. [Online]. Available: [https://doi.org/10.1016/S0957-4174\(02\)00190-2](https://doi.org/10.1016/S0957-4174(02)00190-2)
- [99] D. Leung and J. Romagnoli, "Dynamic probabilistic model-based expert system for fault diagnosis," *Computers & Chemical Engineering*, vol. 24, no. 11, pp. 2473–2492, 2000. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(00\)00610-4](https://doi.org/10.1016/S0098-1354(00)00610-4)
- [100] D. N. C. Phillpotts, "Nonlinear fault detection and diagnosis using kernel based techniques applied to a pilot distillation column," Ph.D. dissertation, University of Pretoria, 2007.
- [101] X. Liu, U. Kruger, T. Littler, L. Xie, and S. Wang, "Moving window kernel PCA for adaptive monitoring of nonlinear processes," *Chemometrics and Intelligent Laboratory Systems*, vol. 96, no. 2, pp. 132–143, 2009. [Online]. Available: <https://doi.org/10.1016/j.chemolab.2009.01.002>
- [102] M. Madakyaru, F. Harrou, and Y. Sun, "Improved data-based fault detection strategy and application to distillation columns," *Process Safety and Environmental Protection*, vol. 107, pp. 22–34, 2017. [Online]. Available: <https://doi.org/10.1016/j.psep.2017.01.017>

- [103] A. Raich and A. Cinar, "Diagnosis of process disturbances by statistical distance and angle measures," *Computers & Chemical Engineering*, vol. 21, no. 6, pp. 661–673, 1997. [Online]. Available: [https://doi.org/10.1016/S0098-1354\(96\)00299-2](https://doi.org/10.1016/S0098-1354(96)00299-2)
- [104] J. Yu, "Local and global principal component analysis for process monitoring," *Journal of Process Control*, vol. 22, no. 7, pp. 1358–1373, 2012. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2012.06.008>
- [105] M. A. Atoui, S. Verron, and A. Kobi, "Fault detection and diagnosis in a bayesian network classifier incorporating probabilistic boundary 1," *IFAC-PapersOnLine*, vol. 48, no. 21, pp. 670–675, 2015. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2015.09.604>
- [106] S. Verron, J. Li, and T. Tiplica, "Fault detection and isolation of faults in a multivariate process with bayesian network," *Journal of Process Control*, vol. 20, no. 8, pp. 902–911, 2010. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2010.06.001>
- [107] M. A. Atoui, S. Verron, and A. Kobi, "Fault detection with conditional gaussian network," *Engineering Applications of Artificial Intelligence*, vol. 45, pp. 473–481, 2015. [Online]. Available: <https://doi.org/10.1016/j.engappai.2015.07.020>
- [108] N. Lu, F. Wang, and F. Gao, "Combination method of principal component and wavelet analysis for multivariate process monitoring and fault diagnosis," *Industrial & Engineering Chemistry Research*, vol. 42, no. 18, pp. 4198–4207, 2003. [Online]. Available: <https://doi.org/10.1021/ie0207313>
- [109] L. Xie, X. Lin, and J. Zeng, "Shrinking principal component analysis for enhanced process monitoring and fault isolation," *Industrial & Engineering Chemistry Research*, vol. 52, no. 49, pp. 17 475–17 486, 2013. [Online]. Available: <https://doi.org/10.1021/ie401030t>
- [110] X. Zhu and R. D. Braatz, "Two-dimensional contribution map for fault identification," *IEEE Control Systems Magazine*, vol. 34, no. 5, pp. 72–77, 2014. [Online]. Available: <https://doi.org/10.1109/MCS.2014.2333295>
- [111] D. Zhou, G. Li, and S. J. Qin, "Total projection to latent structures for process monitoring," *AIChE Journal*, vol. 56, no. 1, pp. 168–178, 2010. [Online]. Available: <https://doi.org/10.1002/aic.11977>
- [112] L. H. Chiang, B. Jiang, X. Zhu, D. Huang, and R. D. Braatz, "Diagnosis of multiple and unknown faults using the causal map and multivariate statistics," *Journal of Process Control*, vol. 28, pp. 27–39, 2015. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2015.02.004>
- [113] E. L. Russell, L. H. Chiang, and R. D. Braatz, "Fault detection in industrial processes using canonical variate analysis and dynamic principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 51, no. 1, pp. 81–93, 2000. [Online]. Available: [https://doi.org/10.1016/S0169-7439\(00\)00058-7](https://doi.org/10.1016/S0169-7439(00)00058-7)

- [114] K. Ghosh, M. Ramteke, and R. Srinivasan, "Optimal variable selection for effective statistical process monitoring," *Computers & Chemical Engineering*, vol. 60, pp. 260–276, 2014. [Online]. Available: <https://doi.org/10.1016/j.compchemeng.2013.09.014>
- [115] B. Wang, X. Yan, Q. Jiang, and Z. Lv, "Generalized Dice's coefficient-based multi-block principal component analysis with Bayesian inference for plant-wide process monitoring," *Journal of Chemometrics*, vol. 29, no. 3, pp. 165–178, 2015. [Online]. Available: <https://doi.org/10.1002/cem.2687>
- [116] Y. Dong and S. J. Qin, "A novel dynamic PCA algorithm for dynamic data modeling and process monitoring," *Journal of Process Control*, vol. 67, pp. 1–11, 2018. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2017.05.002>
- [117] R. Fezai, M. Mansouri, O. Taouali, M. F. Harkat, and N. Bouguila, "Online reduced kernel principal component analysis for process monitoring," *Journal of Process Control*, vol. 61, pp. 1–11, 2018. [Online]. Available: <https://doi.org/10.1016/j.jprocont.2017.10.010>
- [118] S. J. Qin and Y. Zheng, "Quality-relevant and process-relevant fault monitoring with concurrent projection to latent structures," *AIChE Journal*, vol. 59, no. 2, pp. 496–504, 2013. [Online]. Available: <https://doi.org/10.1002/aic.13959>
- [119] E. Russell, L. Chiang, and R. Braatz, "Data-driven methods for fault detection and diagnosis in chemical processes," 2000.
- [120] F.-Y. Lin, Y.-S. Chen, and G.-B. Wang, "Use of Commercial Process Simulator to Mode Transition Control of the Tennessee Eastman Challenge Problem," in *Proceedings of the 2008 AIChE Annual Meeting, Philadelphia, PA, USA*, 2008, pp. 16–21.
- [121] C. Martin-Villalba, A. Urquia, and G. Shao, "Implementations of the Tennessee Eastman Process in Modelica," *IFAC-PapersOnLine*, vol. 51, no. 2, pp. 619–624, 2018. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2018.03.105>
- [122] S. Greyling, H. Marais, G. van Schoor, and K. R. Uren, "Application of Exergy-Based Fault Detection in a Gas-to-Liquids Process Plant," *Entropy*, vol. 21, no. 6, p. 565, 2019. [Online]. Available: <https://doi.org/10.3390/e21060565>
- [123] H. Marais, G. Van Schoor, and K. R. Uren, "Energy-based fault detection for an autothermal reformer," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 353–358, 2016. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2016.07.325>
- [124] L. Fouché, K. Uren, and G. Van Schoor, "Energy-based visualisation of an axial-flow compressor system for the purposes of Fault Detection and Diagnosis," *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 314–319, 2016. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2016.07.311>

- [125] A. N. Stranges, *A History of the Fischer-Tropsch Synthesis in Germany 1926 - 45*. Elsevier, 2007, vol. 163, pp. 1–27.
- [126] A. Steynberg, *Introduction to Fischer-Tropsch Technology*. Elsevier, 2004, vol. 152, ch. 1, pp. 1–63.
- [127] A. Minchener, “Coal-to-oil, gas and chemicals in china,” *IEA Clean Coal Centre*, February 2011.
- [128] Synthetic fuels (methanol to gasoline). [Online]. Available: <https://www.exxonmobilchemical.com/en/catalysts-and-technology-licensing/synthetic-fuels>
- [129] J. Nyári, “Techno-economic feasibility study of a methanol plant using carbon dioxide and hydrogen,” Master’s thesis, KTH School of Industrial Engineering and Management, Stockholm, 2018.
- [130] C. B. Benham, M. S. Bohn, and D. L. Yakobson, “Process for the production of hydrocarbons,” Apr. 1996.
- [131] Dynamic fuels renewable synthetic fuel plant. [Online]. Available: <https://www.chemicals-technology.com/projects/dynamicfuelslouisian/>
- [132] A. De Klerk, *Fischer-Tropsch Refining*, 1st ed. John Wiley & Sons, 2011.
- [133] P. M. Maitlis and A. de Klerk, *Greener Fischer-Tropsch Processes for Fuels and Feedstocks*, 1st ed. John Wiley & Sons, 2013.
- [134] J. R. Rostrup-Nielsen, “New aspects of syngas production and use,” *Catalysis today*, vol. 63, no. 2-4, pp. 159–164, 2000. [Online]. Available: [https://doi.org/10.1016/S0920-5861\(00\)00455-7](https://doi.org/10.1016/S0920-5861(00)00455-7)
- [135] M. W. Smith and D. Shekhawat, “Catalytic partial oxidation,” in *Fuel cells: Technologies for Fuel Processing*. Elsevier, 2011, pp. 73–128.
- [136] B. H. Davis and M. L. Occelli, *Fischer-Tropsch Synthesis, Catalysts and Catalysis*. Elsevier, 2006.
- [137] M. Hillestad, “Modeling the Fischer-Tropsch product distribution and model implementation,” *Chemical Product and Process Modeling*, vol. 10, no. 3, pp. 147–159, 2015. [Online]. Available: <https://doi.org/10.1515/cppm-2014-0031>
- [138] A. A. Al-Yaeshi, A. AlNouss, G. McKay, and T. Al-Ansari, “A simulation study on the effect of CO₂ injection on the performance of the GTL process,” *Computers & Chemical Engineering*, vol. 136, p. 106768, 2020. [Online]. Available: <https://doi.org/10.1016/j.compchemeng.2020.106768>
- [139] F. R. Giordano, M. D. Weir, and W. P. Fox, *A First Course in Mathematical Modeling*, 3rd ed. Thomson Brooks/Cole, 2003.
- [140] A. Al-Matar. Selecting fluid packages (thermodynamic model) for hysys/aspen plus/chemcad process simulators. [Online]. Available: <https://www.researchgate.net/publication/283259774>
- [141] Aspentech’s hysys: Fluid package (thermodynamics) notes. [Online]. Available: <https://smartprocessdesign.com/aspentechs-hysys-fluid-package-thermodynamics-notes/>

- [142] K. E. Vugrin, “On the effect of numerical noise in simulation-based optimization,” Master’s thesis, Virginia Tech, 2003.
- [143] F. Bordoni and A. D’Amico, “Noise in sensors,” *Sensors and Actuators A: Physical*, vol. 21, no. 1-3, pp. 17–24, 1990. [Online]. Available: [https://doi.org/10.1016/0924-4247\(90\)85003-M](https://doi.org/10.1016/0924-4247(90)85003-M)
- [144] K. Aasberg-Petersen, T. S. Christensen, C. S. Nielsen, and I. Dybkjær, “Recent developments in autothermal reforming and pre-reforming for synthesis gas production in GTL applications,” *Fuel Processing Technology*, vol. 83, no. 1-3, pp. 253–261, 2003. [Online]. Available: [https://doi.org/10.1016/S0378-3820\(03\)00073-0](https://doi.org/10.1016/S0378-3820(03)00073-0)
- [145] J. G. Speight, *Gasification of Unconventional Feedstocks*. Gulf Professional Publishing, 2014.
- [146] P. K. Bakkerud, “Update on synthesis gas production for GTL,” *Catalysis Today*, vol. 1, no. 106, pp. 30–33, 2005. [Online]. Available: [10.1016/j.cattod.2005.07.147](https://doi.org/10.1016/j.cattod.2005.07.147)
- [147] E. Iglesia, S. C. Reyes, and S. L. Soled, “Computer-aided design of catalysts,” pp. 199–257, 1993.
- [148] J. McGhee, I. A. Henderson, and A. Baird, “Neural networks applied for the identification and fault diagnosis of process valves and actuators,” *Measurement*, vol. 20, no. 4, pp. 267–275, 1997. [Online]. Available: [https://doi.org/10.1016/S0263-2241\(97\)00043-2](https://doi.org/10.1016/S0263-2241(97)00043-2)
- [149] E. Magnanelli, O. T. Berglihn, and S. Kjelstrup, “Exergy-based performance indicators for industrial practice,” *International Journal of Energy Research*, vol. 42, no. 13, pp. 3989–4007, 2018. [Online]. Available: <https://doi.org/10.1002/er.4123>
- [150] A. Valero, F. Lerch, L. Serra, and J. Royo, “Structural theory and thermoeconomic diagnosis part II: Application to an actual power plant,” *Energy Conversion and Management*, vol. 43, no. 9-12, pp. 1519–1535, 2002. [Online]. Available: [https://doi.org/10.1016/S0196-8904\(02\)00033-X](https://doi.org/10.1016/S0196-8904(02)00033-X)
- [151] A. Zaleta-Aguilar, D. Rodriguez-Alejandro, and V. Rangel-Hernández, “Application of an exergy-based thermo characterization approach to diagnose the operation of a biomass-fueled gasifier,” *Biomass and Bioenergy*, vol. 116, pp. 1–7, 2018. [Online]. Available: <https://doi.org/10.1016/j.biombioe.2018.05.008>
- [152] A. Valero, L. Correas, A. Zaleta, A. Lazzaretto, V. Verda, M. Reini, and V. Rangel, “On the thermoeconomic approach to the diagnosis of energy system malfunctions: Part 1: the TADEUS problem,” *Energy*, vol. 29, no. 12-15, pp. 1875–1887, 2004. [Online]. Available: <https://doi.org/10.1016/j.energy.2004.04.053>
- [153] —, “On the thermoeconomic approach to the diagnosis of energy system malfunctions: Part 2. Malfunction definitions and assessment,” *Energy*, vol. 29, no. 12-15, pp. 1889–1907, 2004. [Online]. Available: <https://doi.org/10.1016/j.energy.2004.03.008>
- [154] M. Shekarchian, F. Zarifi, M. Moghavvemi, F. Motasemi, and T. Mahlia, “Energy, exergy, environmental and economic analysis of industrial fired heaters based on heat recovery and

- preheating techniques,” *Energy Conversion and Management*, vol. 71, pp. 51–61, 2013. [Online]. Available: <https://doi.org/10.1016/j.enconman.2013.03.008>
- [155] W. van Gool, “Thermodynamics of chemical references for exergy analysis,” *Energy Conversion and Management*, vol. 39, no. 16-18, pp. 1719–1728, 1998. [Online]. Available: [https://doi.org/10.1016/S0196-8904\(98\)00089-2](https://doi.org/10.1016/S0196-8904(98)00089-2)
- [156] J. Szargut, “Chemical exergies of the elements,” *Applied Energy*, vol. 32, no. 4, pp. 269–286, 1989.
- [157] W. J. Wepfer and R. A. Gaggioli, “Reference datums for available energy.” ACS Publications, 1980.
- [158] R. Rivero, , and M. Garfias, “Standard chemical exergy of elements updated,” *Energy*, vol. 31, no. 15, pp. 3310–3326, 2006. [Online]. Available: <https://doi.org/10.1016/j.energy.2006.03.020>
- [159] J. R. Muñoz and E. E. Michaelides, “The impact of the model of the environment in exergy analyses,” *ASME, Journal of Energy Resources Technology*, vol. 121, no. 4, pp. 268–277, 1999. [Online]. Available: <https://doi.org/10.1115/1.2795993>
- [160] J. Szargut, “Egzergia: poradnik obliczania i stosowania,” 2007.
- [161] F. Abdollahi-Demneh, M. A. Moosavian, M. R. Omidkhah, and H. Bahmanyar, “Calculating exergy in flowsheeting simulators: A HYSYS implementation,” *Energy*, vol. 36, pp. 5320–5327, 2011. [Online]. Available: <https://doi.org/10.1016/j.energy.2011.06.040>
- [162] J. Bondy and U. Murty, *Graph Theory with Applications*. Elsevier Science Publishing Co., Inc., 1976.
- [163] S. Jouili and S. Tabbone, “Attributed graph matching using local descriptions,” 2009.
- [164] D. R. Wilson and T. R. Martinez, “Improved heterogeneous distance functions,” *Journal of Artificial Intelligence Research*, vol. 6, pp. 1–34, 1997. [Online]. Available: <https://doi.org/10.1613/jair.346>
- [165] S. van Graan, “Graph matching as a means to energy-visualisation of a counter-flow heat exchanger,” Master’s thesis, North-West University Potchefstroom, 2016.
- [166] K. R. Uren, G. Van Schoor, M. Van Eldik, and J. J. De Bruin, “An energy graph-based approach to fault diagnosis of a transcritical CO₂ heat pump,” *Energies*, vol. 13, no. 7, p. 1783, 2020. [Online]. Available: <https://doi.org/10.1016/j.ifacol.2019.09.181>
- [167] H. Abdi, “The eigen-decomposition: Eigenvalues and eigenvectors,” *Encyclopedia of Measurement and Statistics*, pp. 304–308, 2007.
- [168] B. Blakeley, “Autothermal reforming of methane for syngas production in a novel ceramic microchannel reactor,” Ph.D. dissertation, Colorado School of Mines. Arthur Lakes Library, 2015.
- [169] M. Lai, R. Horng, W. Lai, and C. Lee, “Determination of the operating range of CO₂ conversion and syngas production in dry auto-thermal reforming,” *International Journal of Hydrogen Energy*, vol. 38, no. 14, pp. 5705–5712, 2013. [Online]. Available: <https://doi.org/10.1016/j.ijhydene.2013.03.025>

- [170] E. Prosen, K. Pitzer, and F. Rossini, "Heats and free energies of formation of the paraffin hydrocarbons, in the gaseous state, to 1500 K," *Journal of Research of the National Bureau of Standards*, vol. 34, pp. 403–411, 1945.
- [171] J. L. Devore and Farnum, *Applied Statistics for Engineers and Scientists*, 2nd ed., Cengage, Ed. Nelson Education, 2005.

APPENDIX A

CENTRAL COMPOSITE ROTATABLE DESIGN

A.1 INTRODUCTION

Central Composite Rotatable Design (CCRD) is a well-known technique of Experimental Design (ED) that ensures that experiments are identified, defined and executed in a manner that provides the most insight with as little effort as possible. Eventually, a mathematical model is obtained describing the relations between considered factors and observed responses. As mentioned in Section 3.4.4, by adding a carbon dioxide (CO_2) stream as an additional feedstock, the documented feed rates of the methane, steam and oxygen found in literature, would no longer produce syngas of adequate composition and temperature. In order to fine-tune the carbon dioxide (CO_2) flow rate required - instead of randomly altering the flow rate - CCRD was used to evaluate the flow rates to obtain suitable syngas outputs. Figure A.1 shows the fundamental methodology of employing CCRD. The first step is to determine the factors that influenced the responses being investigated. Next, the operating range of the system is defined and assigned. The design matrix is done to outline what the simulation conditions will be. These designed simulations are then executed so that the responses can be recorded. With the responses recorded, the mathematical equation coefficients can be determined by employing least square methods. The obtained mathematical model is finally used to inspect the characteristics of the responses.

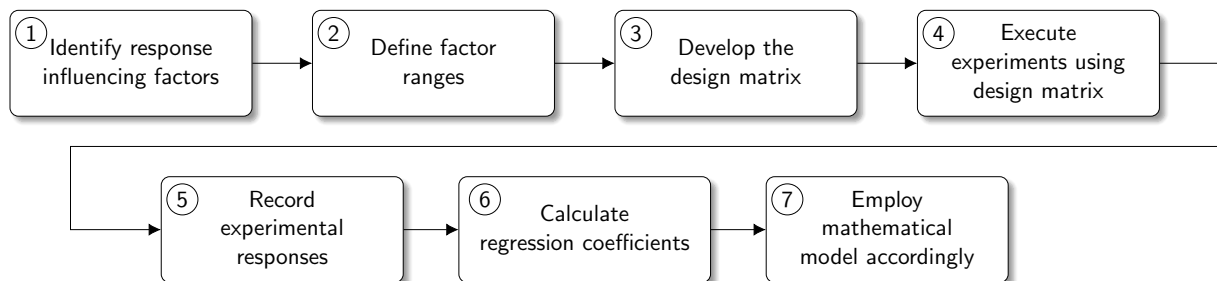


Figure A.1: Generalised methodology of CCRD

A.2 FACTOR IDENTIFICATION

The first important step in Experimental Design is to determine the process variables that will significantly affect the response variables being investigated. For this study, the two response variables were the syngas temperature, denoted as $\Psi_{(T)}$ and the syngas composition, denoted as $\Psi_{(H_2/CO)}$. From literature, it is well-known that the feed ratios of the ATR significantly impact the syngas temperature and composition. The prominent effects that were found documented are:

- The oxygen (O_2) feed stream greatly influences the syngas temperature [24, 168].
- In some plants a carbon dioxide (CO_2) feed stream is used to control the H_2/CO ratio, i.e. the syngas composition [132].

Based on these facts, the influential factors were chosen to be the steam-to-methane ratio (H_2O/CH_4), the oxygen-to-methane ratio (O_2/CH_4), and the carbon dioxide-to-methane ratio (CO_2/CH_4). A summary of the chosen factors and response variables are summarised in Table A.1.

Table A.1: Summary of the factors and response variables used for the CCRD

FACTORS		RESPONSE VARIABLES	
Description	Notation	Description	Notation
Steam-to-methane ratio	H_2O/CH_4	Syngas temperature	$\Psi_{(T)}$
Oxygen-to-methane ratio	O_2/CH_4	Syngas composition	$\Psi_{(H_2/CO)}$
Carbon dioxide-to-methane ratio	CO_2/CH_4		

In order to simulate the considered factors and their effects on the two response variables, the following assumptions were made.

- The methane (CH_4) feed flow rate remained at 8195 kgmole/h.
- The pressure of all the components were kept constant at 3000 kPa.
- The feed temperatures were fixed at:
 - 675 °C for the methane, steam and carbon dioxide streams.
 - 200 °C for the oxygen stream.
- The ATR reaction equations and the comprising attributes were unaltered.

A.3 RANGES OF FACTORS

To utilise the CCRD technique, the investigated factors first need to be assigned viable condition ranges. It is vital that the ranges include the entire spectrum of operating conditions. Here, the minimum and maximum feed ratios were determined from literature [22, 24, 169]. The coded values considered for this CCRD approach were -2 , -1 , 0 , 1 , and 2 . These coded values correspond to range levels described as lowest, low, centre, high and highest. The first actual values were assigned to the low and high levels of

each factor, taking special care to ensure the actual values of lowest and highest would remain within the feasible operating range. The centre value, or g_{x_i} , is calculated by using the actual values of low and high in (A.1). To determine the interval between each level, or t_{x_i} , those same actual values are employed in (A.2). Table A.2 shows the calculated g_{x_i} and t_{x_i} for the three considered factors.

$$g_{x_i} = \frac{\text{high}_{x_i} + \text{low}_{x_i}}{2} \quad (\text{A.1})$$

$$t_{x_i} = \frac{\text{high}_{x_i} - \text{low}_{x_i}}{2} \quad (\text{A.2})$$

Table A.2: The calculated g_{x_i} and t_{x_i} values for the three factors

Factor	Symbol	low $_{x_i}$	high $_{x_i}$	g_{x_i}	t_{x_i}
H ₂ O/CH ₄	x_1	0.925	1.975	1.450	0.525
O ₂ /CH ₄	x_2	0.575	0.725	0.650	0.075
CO ₂ /CH ₄	x_3	0.05375	0.15125	0.10250	0.04875

To complete the assignments of the actual values of the factor levels, the calculated g_{x_i} and t_{x_i} were incorporated in (A.3). The factors' actual and corresponding coded values are summarised in Table A.3.

$$\text{actual value} = g_{x_i} + (\text{coded value} \cdot t_{x_i}) \quad (\text{A.3})$$

Table A.3: The actual and coded values of the three factors

Factor		Actual value				
		lowest $_{x_i}$	low $_{x_i}$	centre $_{x_i}$	high $_{x_i}$	highest $_{x_i}$
		-2	-1	0	1	2
H ₂ O/CH ₄	x_1	0.400	0.925	1.450	1.975	2.500
O ₂ /CH ₄	x_2	0.500	0.575	0.650	0.725	0.800
CO ₂ /CH ₄	x_3	0.00500	0.05375	0.10250	0.15125	0.20000

A.4 DESIGN MATRIX

With the factors having been assigned actual and coded values, a design matrix could be developed. A design matrix consists of a specific sequence of the coded factors as defined for a CCRD approach with three variables ($k = 3$), called the Yates standard. The design matrix, the actual values used within HYSYS[®] as well as the responses recorded for each simulation run, is tabulated in Table A.4.

A.5 MATHEMATICAL EQUATIONS

With the simulations run and the various responses recorded, it was possible to develop the mathematical equations that describe the response variable in terms of the H₂O/CH₄, O₂/CH₄ and CO₂/CH₄ factors. For

this approach the mathematical model consists of main effect terms of each factor, quadratic terms of each of the factors and first order interaction terms for each paired combination of factors; the general form is shown in (A.4).

$$\Psi = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_1^2 + \beta_5x_2^2 + \beta_6x_3^2 + \beta_7x_1x_2 + \beta_8x_1x_3 + \beta_9x_2x_3 \quad (\text{A.4})$$

The compressed form of the equation is given in (A.5) where $\bar{\Psi}$ represents the matrix of the recorded response values, $\bar{\beta}$ represents the unknown coefficient matrix, $\bar{X}$ the matrix containing the independent factors, given in (A.6), and ϵ the error matrix.

$$\bar{\Psi} = \bar{\beta}\bar{X} + \epsilon \quad (\text{A.5})$$

$$\bar{X} = \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 4 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -2 & 0 & 0 & 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 & 4 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (\text{A.6})$$

In order to calculate the unknown coefficients $\bar{\beta}$ of the syngas temperature and composition equations, the `lskov` function within MATLAB[®] was utilised. By executing $\bar{\beta} = \text{lskov}(\bar{X}, \bar{\Psi})$, the function determines the ordinary least squares (OLS) solution to the described equation in (A.5). After computing the coefficients for the syngas temperature and composition separately, the mathematical equations shown in (A.7) and (A.8) were obtained.

Table A.4: The design matrix along with the two response variables' simulation values

Simulation run	Coded value			Actual value				Simulation results				
	x_1	x_2	x_3	x_1	x_2	x_3	CH ₄	H ₂ O	O ₂	CO ₂	$\Psi_{(T)}$	$\Psi_{(H_2/CO)}$
1	-1	-1	-1	0.925	0.575	0.0540	8195.0	7580.4	4712.1	442.5	1087	2.321
2	1	-1	-1	1.975	0.575	0.0540	8195.0	16185.1	4712.1	442.5	1014	3.036
3	-1	1	-1	0.925	0.725	0.0540	8195.0	7580.4	5941.4	442.5	1516	1.900
4	1	1	-1	1.975	0.725	0.0540	8195.0	16185.1	5941.4	442.5	1341	2.365
5	-1	-1	1	0.925	0.575	0.1510	8195.0	7580.4	4712.1	1237.4	1065	2.118
6	1	-1	1	1.975	0.575	0.1510	8195.0	16185.1	4712.1	1237.4	1001	2.793
7	-1	1	1	0.925	0.725	0.1510	8195.0	7580.4	5941.4	1237.4	1478	1.713
8	1	1	1	1.975	0.725	0.1510	8195.0	16185.1	5941.4	1237.4	1317	2.156
9	-2	0	0	0.400	0.650	0.1025	8195.0	3278.0	5326.8	840.0	1372	1.749
10	2	0	0	2.500	0.650	0.1025	8195.0	20487.5	5326.8	840.0	1122	2.843
11	0	-2	0	1.450	0.500	0.1025	8195.0	11882.8	4097.5	840.0	930.3	2.852
12	0	2	0	1.450	0.800	0.1025	8195.0	11882.8	6556.0	840.0	1580	1.840
13	0	0	-2	1.450	0.650	0.0050	8195.0	11882.8	5326.8	41.0	1241	2.475
14	0	0	2	1.450	0.650	0.2000	8195.0	11882.8	5326.8	1639.0	1191	2.061
15	0	0	0	1.450	0.650	0.1025	8195.0	11882.8	5326.8	840.0	1215	2.248
16	0	0	0	1.450	0.650	0.1025	8195.0	11882.8	5326.8	840.0	1215	2.248
17	0	0	0	1.450	0.650	0.1025	8195.0	11882.8	5326.8	840.0	1215	2.248
18	0	0	0	1.450	0.650	0.1025	8195.0	11882.8	5326.8	840.0	1215	2.248
19	0	0	0	1.450	0.650	0.1025	8195.0	11882.8	5326.8	840.0	1215	2.248
20	0	0	0	1.450	0.650	0.1025	8195.0	11882.8	5326.8	840.0	1215	2.248

$$\Psi_{(T)} = 1213.9 - 60.8x_1 + 174.03x_2 - 12.3x_3 + 7.5x_1x_2 + 9.5x_2x_3 - 0.29x_1x_3 - 24.9x_1^2 + 2.9x_2^2 - 3.4x_3^2 \quad (\text{A.7})$$

$$\Psi_{(\text{H}_2/\text{CO})} = 2.25 + 0.28x_1 - 0.26x_2 - 0.11x_3 + 0.01x_1x_2 + 0.02x_2x_3 + 0.01x_1x_3 - 0.06x_1^2 - 0.01x_2^2 + 0.01x_3^2 \quad (\text{A.8})$$

A.6 RESPONSE EVALUATION

By employing these mathematical equations, the factors and their different effects on the responses can be evaluated in a useful and structured manner. In order to visualise these effects, response surface (RS) plots are usually developed and assessed. The response surface is plotted by calculating the response equation output by sweeping two of the three factors from -2 to 2 , whilst the third factor is kept at a constant value, usually 0 . To determine the necessary feed ratios for the syngas temperature and composition from the RS plots, the following steps were done:

1. The first step is to assess the syngas temperature response $\Psi_{(T)}$ plot. Seeing as the O_2/CH_4 is known to affect the temperature the most, the response was evaluated in terms of $\text{H}_2\text{O}/\text{CH}_4$ and O_2/CH_4 , while keeping $\text{CO}_2/\text{CH}_4 = 0$. The RS plot depicting these effects is shown in Figure A.2a. From literature, the $\text{H}_2\text{O}/\text{CH}_4$ ratio is seen to range from $0.6 - 0.7$ []. Therefore, to simplify the evaluation process, the actual value was fixed at 0.6625 , i.e. coded value of $\text{H}_2\text{O}/\text{CH}_4 = -1.5$. The coded value is easily obtained by rewriting (A.3) to the form given in (A.9)

$$\text{coded value} = \frac{\text{actual value} - g_{x_i}}{t_{x_i}} \quad (\text{A.9})$$

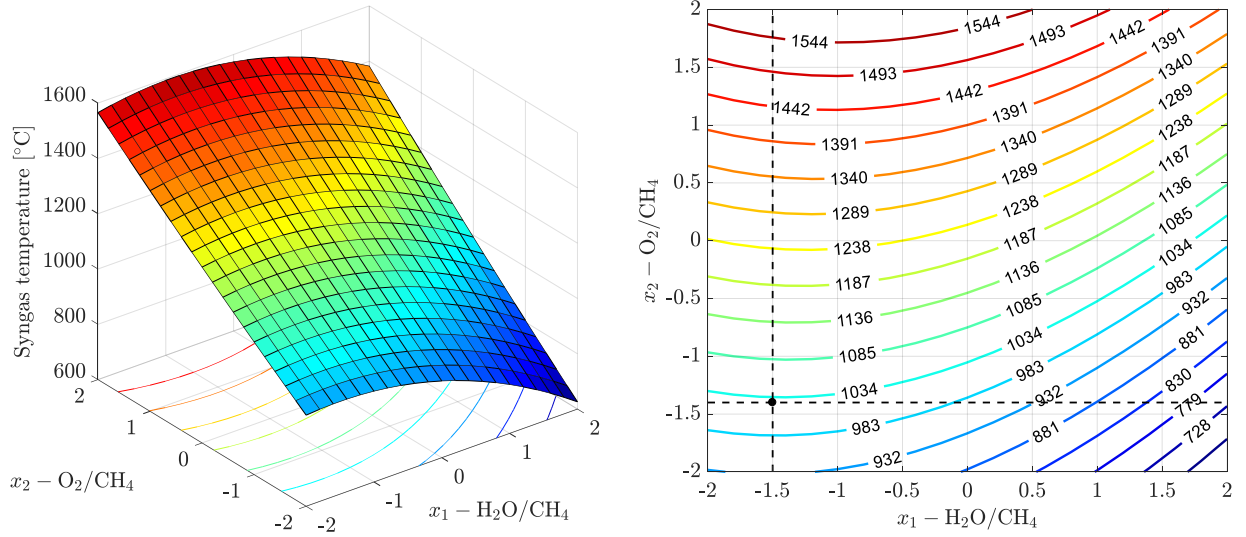
Viewing the RS plot in Figure A.2a, it is seen that for a specific $\text{H}_2\text{O}/\text{CH}_4$ ratio, the temperature will increase as the O_2/CH_4 ratio increases. In order to obtain a temperature of $\approx 1030^\circ\text{C}$, a lower end O_2/CH_4 ratio would be required along with the chosen $\text{H}_2\text{O}/\text{CH}_4$. Visual inspection would suggest a coded value $\text{O}_2/\text{CH}_4 = -1.4$ would deliver an appropriate temperature. By substituting these coded values in (A.7), the expected temperature was $\approx 1027^\circ\text{C}$.

2. As the CO_2/CH_4 ratio is used to control the syngas composition $\Psi_{(\text{H}_2/\text{CO})}$, the impact of the $\text{H}_2\text{O}/\text{CH}_4$ and CO_2/CH_4 ratios when $\text{O}_2/\text{CH}_4 = -1.4$, needs to be assessed. To do this, the RS plot in Figure A.2b was examined. If the $\text{H}_2\text{O}/\text{CH}_4$ coded value were to remain -1.5 , a coded value of 0.1 for the CO_2/CH_4 ratio would produce an adequate syngas composition of ≈ 2.05 .

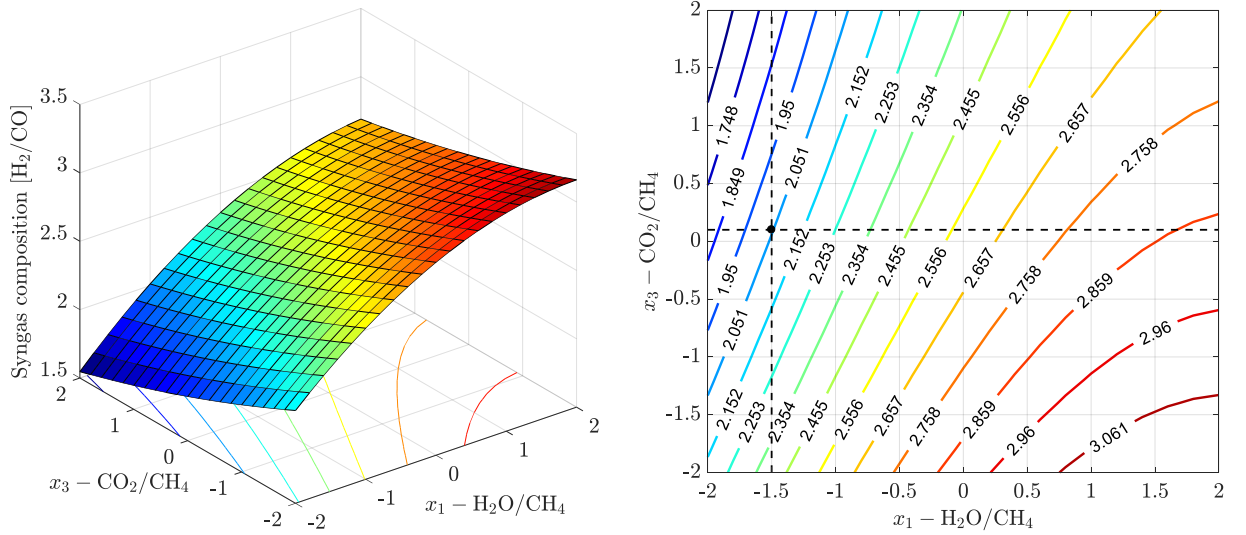
These coded values for the three ratios were converted back to actual values by employing (A.3) accordingly. These ratios were then further transformed to actual feed flow rates per stream and are summarised in Table A.5. The HYSYS® model was simulated using these assumed values in order to determine whether the syngas temperature and composition were what was expected. Small deviations were seen in the syngas outputs; the syngas temperature was slightly higher than the anticipated 1027°C , and the composition was slightly above 2.05 . This prompted minor alterations to the ratios to obtain the required syngas characteristics. The eventual ratios, which are sufficiently close to the assumed values, are also summarised in Table A.5.

Table A.5: Assumed and eventual values for the chosen feed ratios

RATIO					Component	MOLAR FLOW RATE	
Assumed		Eventual		Assumed		Eventual	
Coded	Actual	Coded	Actual				
H ₂ O/CH ₄	−1.50	0.6625	−1.50	0.6625	H ₂ O	5429.2	5429.2
O ₂ /CH ₄	−1.40	0.5450	−1.45	0.5450	O ₂	4466.3	4434.9
CO ₂ /CH ₄	0.10	0.1074	0.30	0.1074	CO ₂	879.9	959.4



(a) The effect of x_1 and x_2 on the syngas temperature with $x_3 = 0$



(b) The effect of x_1 and x_3 on the syngas composition with $x_2 = -1.4$

Figure A.2: The response surface plots for the effects on (a) temperature by x_1 and x_2 , (b) composition by x_1 and x_3

APPENDIX B

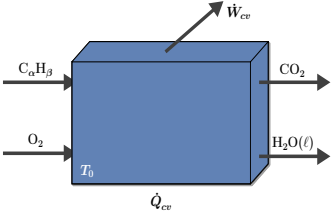
STANDARD CHEMICAL EXERGY CALCULATIONS

As stated in Section 4.3.4.2, the molar chemical exergy of substance i is computed by making use of:

$$b_{ch} = \sum x_{(i)} b_{ch(i)}^0 \quad (\text{B.1})$$

where $x_{(i)}$ is the mole fraction and $b_{ch(i)}^0$ the standard chemical exergy of substance i . For most substances, the standard chemical exergy ($b_{ch(i)}^0$) is readily available. Unfortunately, for some of the considered hydrocarbons, no such values were found documented. Dincer et al. [25], however, illustrate that for any unknown substance, the standard chemical exergy can be calculated by utilising an appropriate reaction equation. To determine the unavailable hydrocarbons, the combustion with oxygen depicted in Table B.1 (a) and of which the equation is given in (B.2), is used.

Table B.1: (a) Hypothetical chamber showing fuel conversion and (b) the corresponding combustion equation

(a) Diagram	(b) Equation
	$\text{C}_\alpha \text{H}_\beta + \left(\alpha + \frac{\beta}{4} \right) \text{O}_2 \rightarrow \alpha \text{CO}_2 + \frac{\beta}{2} \text{H}_2\text{O}(\ell) \quad (\text{B.2})$

In (B.2), the α and β coefficients will correspond to the hydrocarbon being considered. By modifying this equation considerably, as comprehensively detailed in [25], the following equation can be derived:

$$b_{ch(i)}^0 = \left[\bar{g}_{(i)}^0 + \left(\alpha + \frac{\beta}{4} \right) \bar{g}_{(\text{O}_2)}^0 - \alpha \bar{g}_{(\text{CO}_2)}^0 - \frac{\beta}{2} \bar{g}_{(\text{H}_2\text{O}(\ell))}^0 \right] + \alpha b_{ch(\text{CO}_2)}^0 + \frac{\beta}{2} b_{ch(\text{H}_2\text{O}(\ell))}^0 - \left(\alpha + \frac{\beta}{4} \right) b_{ch(\text{O}_2)}^0 \quad (\text{B.3})$$

where $\bar{g}^0$ refers to the Gibbs function of formation and b_{ch}^0 the known standard chemical exergy of each component shown in brackets. Finding Gibbs function of formation values for the hydrocarbons proved difficult. Eventually, the free energies of formations (F_f^0) documented in [170] were converted by making

use of the conversion rate $1 \text{ kcal/mole} \Leftrightarrow 4183 \text{ kJ/kgmole}$. Table B.2 summarises the values used within (B.3).

Table B.2: Gibbs of formation and standard chemical exergy values used to calculate the hydrocarbons

$\bar{g}^0 \text{ [kJ/kgmole]}$			$b_{ch}^0 \text{ [kJ/kgmole]}$		
$\bar{g}_{(O_2)}^0$	$\bar{g}_{(CO_2)}^0$	$\bar{g}_{(H_2O(\ell))}^0$	$b_{ch(CO_2)}^0$	$b_{ch(H_2O(\ell))}^0$	$b_{ch(O_2)}^0$
0	-394360	-237180	19480	900	3970

To validate that the computed standard chemical exergy quantities were adequate, the values were compared to known hydrocarbon values found in literature. The hydrocarbons considered were CH_4 - C_5H_{12} with the results tabulated in Table B.3.

Table B.3: Comparison between known and calculated standard chemical exergy

Substance	Formula	$b_{ch}^0 \text{ [kJ/kgmole]}$		Difference [%]
		Tabulated [†]	Calculated	
Methane	CH_4	831200	831275	0.009
Ethane	C_2H_6	1495000	1495144	0.010
Propane	C_3H_8	2152800	2150505	0.107
n-Butane	C_4H_{10}	2804200	2804251	0.002
n-Pentane	C_5H_{12}	3461300	3457721	0.103

[†] Values obtained from Table B.7

Based on the marginal differences seen, the method for calculating the hydrocarbon exergies was deemed appropriate. Table B.4 and Table B.5 shows the substances' corresponding formation energy and the obtained standard chemical exergy. Table B.6 summarises the standard chemical exergy values for the remaining components included in the simulation.

Table B.4: Gibbs function of formation and calculated standard chemical exergy CH_4 - $C_{11}H_{24}$

Substance	Formula	$\Delta F_f^0 \text{ [kcal/mole]}^\ddagger$	Calculated	
			$\bar{g}^0 \text{ [kJ/kgmole]}$	$b_{ch}^0 \text{ [kJ/kgmole]}$
Methane	CH_4	-12.140	-50785.262	831275
Ethane	C_2H_6	-7.860	-32880.738	1495144
Propane	C_3H_8	-5.614	-23485.046	2150505
n-Butane	C_4H_{10}	-3.754	-15704.108	2804251
n-Pentane	C_5H_{12}	-1.960	-8199.268	3457721
n-Hexane	C_6H_{14}	0.050	209.165	4112094
n-Heptane	C_7H_{16}	2.090	8743.097	4766593
n-Octane (g)	$C_8H_{18}(g)$	4.140	17318.862	5421134
n-Octane (ℓ)	$C_8H_{18}(\ell)$	-	6600.000	5410415
n-Nonane	C_9H_{20}	6.180	25852.794	6075633
n-Decane	$C_{10}H_{22}$	8.230	34428.559	6730174
n-Undecane	$C_{11}H_{24}$	10.280	43004.324	7384714

[‡] Values obtained from Figure B.1 and Figure B.2 in 298.16 K column

Table B.5: Gibbs function of formation and calculated standard chemical exergy of $C_{12}H_{26}$ - $C_{30}H_{62}$

Substance	Formula	ΔF_f^0 [kcal/mole] [‡]	Calculated	
			$\bar{g}^0$ [kJ/kgmole]	b_{ch}^0 [kJ/kgmole]
n-Dodecane	$C_{12}H_{26}$	12.330	51580.089	8039255
n-Tridecane	$C_{13}H_{28}$	14.370	60114.021	8693754
n-Tetradecane	$C_{14}H_{30}$	16.420	68689.786	9348295
n-Pentadecane	$C_{15}H_{32}$	18.470	77265.551	10002836
n-Hexadecane	$C_{16}H_{34}$	20.520	85841.316	10657376
n-Heptadecane	$C_{17}H_{36}$	22.560	94375.248	11311875
n-Octadecane	$C_{18}H_{38}$	24.610	102951.013	11966416
n-Nonadecane	$C_{19}H_{40}$	26.660	111526.778	12620957
n-Eicosane	$C_{20}H_{42}$	28.710	120102.543	13275498
Lumped	$C_{30}H_{62}$	47.142	197209.129	19812254

[‡] Values obtained from Figure B.1 and Figure B.2 in 298.16 K column

Table B.6: Standard chemical exergy of the other substances included in simulation [25]

Substance	Formula	b_{ch}^0 [kJ/kgmole] [†]
Carbon monoxide	CO	274710
Carbon dioxide	CO ₂	19480
Hydrogen	H ₂	236090
Water vapour	H ₂ O(g)	9500
Water	H ₂ O(ℓ)	900
Oxygen	O ₂	3970

[†] Values obtained from Table B.7

TABLE 2.—Values of the free energies of formation of the paraffin hydrocarbons in the gaseous state to 1,500 °K

Compound (gas)	Temperature in °K														
	0	298.16	300	400	500	600	700	800	900	1,000	1,100	1,200	1,300	1,400	1,500
	ΔF_f° in kcal/mole														
Methane.....	-15.987	-12.140	-12.105	-10.048	-7.841	-5.49	-3.05	-0.55	2.01	4.61	7.22	9.85	12.50	15.14	17.79
Ethane.....	-16.517	-7.860	-7.785	-3.447	1.168	5.97	10.90	15.92	21.00	26.13	31.28	36.45	41.62	46.79	51.99
Propane.....	-19.482	-5.614	-5.541	1.191	8.230	15.50	22.93	30.45	38.05	45.68	53.34	61.01	68.68	76.35	84.08
<i>n</i> -Butane.....	-23.332	-3.754	-3.595	5.435	14.868	24.59	34.50	44.50	54.59	64.71	74.87	85.05	95.25	105.42	115.61
2-Methylpropane.....	-24.602	-4.296	-4.127	5.273	15.088	25.18	35.44	45.82	56.29	66.79	77.32	87.86	98.39	108.90	119.44
<i>n</i> -Pentane.....	-27.27	-1.96	-1.76	9.66	21.57	33.79	46.22	58.77	71.41	84.10	96.81	109.55	122.30	135.01	147.75
2-Methylbutane.....	-28.66	-3.50	-3.38	8.26	20.29	32.66	45.19	57.84	70.62	83.41	96.23	109.07	121.89	134.68	147.53
2,2-Dimethylpropane.....	-31.30	-3.64	-3.42	8.99	21.89	35.08	48.46	61.93	75.49	89.08	102.68	116.27	129.88	143.44	157.00
<i>n</i> -Hexane.....	-30.98	0.05	0.30	14.11	28.46	43.18	58.13	73.21	88.40	103.64	118.90	134.19	149.48	164.75	180.04
2-Methylpentane.....	-32.03	-0.96	-0.71	13.25	27.79	42.72	57.88	73.15	88.51	103.86	-----	-----	-----	-----	-----
3-Methylpentane.....	-31.36	-0.29	-0.04	13.94	28.50	43.45	58.62	73.89	89.23	104.53	-----	-----	-----	-----	-----
2,2-Dimethylbutane.....	-34.52	-2.35	-2.09	12.30	27.30	42.69	58.29	73.97	89.71	105.37	-----	-----	-----	-----	-----
2,3-Dimethylbutane.....	-32.82	-0.73	-0.47	13.81	28.67	43.92	59.38	74.94	90.57	106.18	-----	-----	-----	-----	-----
<i>n</i> -Heptane.....	-34.65	2.09	2.38	18.58	35.39	52.60	70.08	87.70	105.43	123.22	141.02	158.85	176.69	194.50	212.38
2-Methylhexane.....	-35.77	0.98	1.27	17.61	34.60	52.01	69.67	87.47	105.33	123.18	-----	-----	-----	-----	-----
3-Methylhexane.....	-34.96	1.10	1.39	17.77	34.40	51.68	69.21	86.87	104.59	122.29	-----	-----	-----	-----	-----
3-Ethylpentane.....	-34.10	2.59	2.88	19.36	36.52	54.15	72.06	90.09	108.19	126.25	-----	-----	-----	-----	-----
2,2-Dimethylpentane.....	-38.00	0.09	0.39	17.33	34.95	53.01	71.33	89.73	108.18	126.55	-----	-----	-----	-----	-----
2,3-Dimethylpentane.....	-36.29	0.16	0.45	16.83	33.90	51.43	69.24	87.17	105.18	123.16	-----	-----	-----	-----	-----
2,4-Dimethylpentane.....	-36.98	0.72	1.02	17.81	35.28	53.19	71.37	89.66	107.99	126.27	-----	-----	-----	-----	-----
3,3-Dimethylpentane.....	-36.92	0.63	0.93	17.64	35.02	52.84	70.92	89.10	107.34	125.53	-----	-----	-----	-----	-----
2,2,3-Trimethylbutane.....	-37.57	0.76	1.06	18.06	35.75	53.87	72.24	90.72	109.23	127.68	-----	-----	-----	-----	-----
<i>n</i> -Octane.....	-38.33	4.14	4.47	23.06	42.32	62.03	82.03	102.17	122.45	142.79	163.14	183.51	203.90	224.26	244.71
2-Methylheptane.....	-39.42	3.06	3.39	22.14	41.58	61.47	81.63	101.88	122.27	142.83	-----	-----	-----	-----	-----
3-Methylheptane.....	-38.64	3.29	3.62	22.21	41.50	61.27	81.31	101.44	121.68	142.07	-----	-----	-----	-----	-----
4-Methylheptane.....	-38.43	4.00	4.34	23.12	42.62	62.61	82.88	103.23	123.68	144.26	-----	-----	-----	-----	-----
3-Ethylhexane.....	-37.71	3.95	4.28	22.94	42.39	62.37	82.64	102.98	123.42	143.98	-----	-----	-----	-----	-----
2,2-Dimethylhexane.....	-41.23	2.56	2.91	22.18	42.20	62.71	83.48	104.30	125.21	146.27	-----	-----	-----	-----	-----

Figure B.1: Free energies of formation for various hydrocarbons for temperature range 0 - 1500 K [170]

2,3-Dimethylhexane	-38.76	4.23	4.57	23.51	43.16	63.27	83.64	104.09	124.63	145.30					
2,4-Dimethylhexane	-39.74	2.80	3.14	22.05	41.72	61.89	82.32	102.84	123.43	144.07					
2,5-Dimethylhexane	-40.61	2.50	2.84	21.92	41.75	62.08	82.69	103.41	124.20	144.99					
3,3-Dimethylhexane	-39.90	3.17	3.51	22.63	42.50	62.85	83.47	104.16	124.93	145.76					
3,4-Dimethylhexane	-38.52	4.97	5.31	24.42	44.27	64.57	85.12	105.75	126.49	147.35					
2-Methyl-3-ethylpentane	-37.96	5.08	5.43	24.47	44.24	64.47	84.96	105.52	126.21	147.09					
3-Methyl-3-ethylpentane	-38.68	4.76	5.11	24.32	44.30	64.77	85.48	106.22	127.05	148.01					
2,2,3-Trimethylpentane	-39.77	4.09	4.44	23.82	43.98	64.63	85.54	106.52	127.57	148.68					
2,2,4-Trimethylpentane	-40.73	3.13	3.48	22.86	43.02	63.67	84.58	105.56	126.61	147.72					
2,3,3-Trimethylpentane	-39.01	4.52	4.86	24.09	44.08	64.55	85.30	106.10	126.98	147.90					
2,3,4-Trimethylpentane	-39.12	4.32	4.67	23.89	43.89	64.37	85.11	105.94	126.84	147.80					
2,2,3,3-Tetramethylbutane	-41.09	4.88	5.24	25.34	46.19	67.47	88.97	110.52	132.19	154.05					
<i>n</i> -Nonane	-42.00	6.18	6.56	27.54	49.25	71.46	93.98	116.66	139.48	162.37	184.20	208.18	231.12	254.03	277.04
<i>n</i> -Decane	-45.67	8.23	8.65	32.01	56.18	80.89	105.93	131.14	156.51	181.95	205.26	232.85	258.33	283.79	309.37
<i>n</i> -Undecane	-49.34	10.28	10.74	36.49	63.11	90.31	117.88	145.63	173.54	201.53	226.32	257.52	285.54	313.55	341.70
<i>n</i> -Dodecane	-53.02	12.33	12.83	40.97	70.04	99.74	129.83	160.11	190.57	221.11	247.38	282.19	312.76	343.32	374.03
<i>n</i> -Tridecane	-56.69	14.37	14.92	45.45	76.97	109.17	141.78	174.59	207.60	240.69	268.45	306.86	339.97	373.08	406.36
<i>n</i> -Tetradecane	-60.36	16.42	17.01	49.93	83.90	118.60	153.73	189.08	224.63	260.26	289.51	331.52	367.19	402.85	438.69
<i>n</i> -Pentadecane	-64.04	18.47	19.10	54.41	90.83	128.03	165.68	203.56	241.66	279.84	310.57	356.19	394.40	432.61	471.02
<i>n</i> -Hexadecane	-67.71	20.52	21.19	58.89	97.76	137.46	177.63	218.04	258.69	299.42	331.63	380.86	421.62	462.38	503.36
<i>n</i> -Heptadecane	-71.38	22.56	23.27	63.36	104.69	146.88	189.58	232.53	275.72	319.00	352.69	405.53	448.83	492.14	535.69
<i>n</i> -Octadecane	-75.06	24.61	25.36	67.84	111.62	156.31	201.53	247.01	292.75	338.58	373.75	430.20	476.05	521.90	568.02
<i>n</i> -Nonadecane	-78.73	26.66	27.45	72.32	118.55	165.74	213.48	261.49	309.78	358.16	394.81	454.86	503.26	551.67	600.35
<i>n</i> -Eicosane	-82.40	28.71	29.54	76.80	125.48	175.17	225.43	275.98	326.81	377.74	415.87	479.53	530.48	581.43	632.68
Δ per CH ₂	-3.673	2.048	2.090	4.479	6.931	9.428	11.951	14.484	17.029	19.579	21.061	24.668	27.215	29.764	32.331

Heats and Free Energies of Formation

Figure B.2: Free energies of formation for various hydrocarbons for temperature range 0 - 1500 K (continued) [170]

Table B.7: Gibbs function of formation ($\bar{g}^0$) for some common substances

Enthalpy of formation, Gibbs function of formation, and absolute entropy at 25°C, 1 atm

Substance	Formula	$\bar{h}_f^0$ kJ/kmol	$\bar{g}_f^0$ kJ/kmol	$\bar{s}^0$ kJ/kmol·K
Carbon	C(s)	0	0	5.74
Hydrogen	H ₂ (g)	0	0	130.68
Nitrogen	N ₂ (g)	0	0	191.61
Oxygen	O ₂ (g)	0	0	205.04
Carbon monoxide	CO(g)	−110,530	−137,150	197.65
Carbon dioxide	CO ₂ (g)	−393,520	−394,360	213.80
Water vapor	H ₂ O(g)	−241,820	−228,590	188.83
Water	H ₂ O(l)	−285,830	−237,180	69.92
Hydrogen peroxide	H ₂ O ₂ (g)	−136,310	−105,600	232.63
Ammonia	NH ₃ (g)	−46,190	−16,590	192.33
Methane	CH ₄ (g)	−74,850	−50,790	186.16
Acetylene	C ₂ H ₂ (g)	+226,730	+209,170	200.85
Ethylene	C ₂ H ₄ (g)	+52,280	+68,120	219.83
Ethane	C ₂ H ₆ (g)	−84,680	−32,890	229.49
Propylene	C ₃ H ₆ (g)	+20,410	+62,720	266.94
Propane	C ₃ H ₈ (g)	−103,850	−23,490	269.91
<i>n</i> -Butane	C ₄ H ₁₀ (g)	−126,150	−15,710	310.12
<i>n</i> -Octane	C ₈ H ₁₈ (g)	−208,450	+16,530	466.73
<i>n</i> -Octane	C ₈ H ₁₈ (l)	−249,950	+6,610	360.79
<i>n</i> -Dodecane	C ₁₂ H ₂₆ (g)	−291,010	+50,150	622.83
Benzene	C ₆ H ₆ (g)	+82,930	+129,660	269.20
Methyl alcohol	CH ₃ OH(g)	−200,670	−162,000	239.70
Methyl alcohol	CH ₃ OH(l)	−238,660	−166,360	126.80
Ethyl alcohol	C ₂ H ₅ OH(g)	−235,310	−168,570	282.59
Ethyl alcohol	C ₂ H ₅ OH(l)	−277,690	−174,890	160.70
Oxygen	O(g)	+249,190	+231,770	161.06
Hydrogen	H(g)	+218,000	+203,290	114.72
Nitrogen	N(g)	+472,650	+455,510	153.30
Hydroxyl	OH(g)	+39,460	+34,280	183.70

Source: From JANAF, *Thermochemical Tables* (Midland, MI: Dow Chemical Co., 1971); *Selected Values of Chemical Thermodynamic Properties*, NBS Technical Note 270-3, 1968; and *API Research Project 44* (Carnegie Press, 1953).

APPENDIX C

HYSYS[®] USER VARIABLES

The appendix gives some details on the *user variables*, used in Chapter 4, that were developed within HYSYS[®] to automatically compute the physical and chemical exergy of the GTL streams. In the first section, the *user variables*' set-up and selections are documented. The second section shows the actual source code as it was implemented throughout the GTL simulation model.

C.1 SET-UP

In order to create a *user variable* in HYSYS[®], one starts off by opening the considered *Material Stream*. Here from *Worksheet* >> *User Variables* >> *Create a New User Variable*. It is imperative that the *user variable* is given an appropriate *Name*. Table C.1 summarises the presented options within the user variable window, and the corresponding selection made for each.

Table C.1: User variable option setup

OPTION	Type	Dimension	Units	Macro	Activation
SELECTION	Real	Scalar	Multiplier	☒ PostExecute()	User Enable

C.2 VBA CODE

With the *user variable* options set-up accordingly, the Visual Basic for Applications (VBA) code could be implemented. Figure C.1 shows the physical exergy calculation, and Figure C.2 the chemical exergy. As multiple phases exist in some of the streams, the chemical exergy was computed per phase present. The code provided in Figure C.2 calculates the vapour chemical exergy. To calculate the liquid and aqueous phases, the same code with two minor tweaks is applied. The first alteration made, indicated by ①, is done to extract the phase-specific molar flow of component *i*. Therefore, depending on the phase being considered, the `mi` value is set by the corresponding `Stream.MolarFlows.Value` command. The second modification ② is to ensure that the correct phase's standard chemical exergy, saved as *user properties*, is called. For the vapour phase the applicable standard chemical exergy is stored as a *user property* named `bch`. The liquid and aqueous phases use the values stored in *user property* `bchl`.

```

'SUBROUTINE THAT CALCULATES THE PHYSICAL EXERGY -----

Sub PostExecute()
    On Error GoTo ErrorHandler

    Dim Stream As Object
    Dim bph,h,h0,s,s0,T0,P0 As Double

    T0 = 25.0 'The SRE temperature (C)
    P0 = 101.325 'The SRE pressure (kPa)

    Set Stream = ActiveObject.DuplicateFluid

    If(Stream.MolarFlow.IsKnown And Stream.MolarFractions.IsKnown(0)
    And Stream.VapourFraction.IsKnown And Stream.Pressure.IsKnown
    And
    T0 <> -32767 And P0 <> -32767) Then
        h = Stream.MolarEnthalpy.GetValue("kJ/kgmole")
        s = Stream.MolarEntropy.GetValue("kJ/kgmole-C")
        Stream.Temperature.SetValue(T0,"C")
        Stream.Pressure.SetValue(P0,"kPa")
        Stream.TPFlash()
        h0 = Stream.MolarEnthalpy.GetValue("kJ/kgmole")
        s0 = Stream.MolarEntropy.GetValue("kJ/kgmole-C")
        bph = (h-h0)-(T0+273.15)*(s-s0)
        bph = bph*Stream.MolarFlow.GetValue("kgmole/h")
        ActiveVariableWrapper.Variable.SetValue(bph)

    End If
ErrorHandler:
End Sub

```

Figure C.1: The *user variable* created to calculate the physical exergy

```

'SUBROUTINE THAT CALCULATES THE CHEMICAL EXERGY FOR A STREAM -----

Sub PostExecute()
    On Error GoTo ErrorHandler

    Dim Stream As Object
    Dim Components As HYSYS.Components
    Dim Component As HYSYS.Component

    'Instances used in calculating the formulae
    Dim mi, ratioi, moleFracPhasei, bchPhase As Double
    Dim mi_T, moleFraci_T, bch_T As Double

    Set Stream = ActiveObject.DuplicateFluid

    If (Stream.Pressure.IsKnown And Stream.VapourFraction.IsKnown And
        Stream.MolarFlow.IsKnown And Stream.MolarFractions.IsKnown(0)) Then

        Set Components = Stream.Components
        bchPhase = 0

        For i=0 To Components.Count-1

            ratioi = 0

            If (Stream.MolarFlows.Values(i) > 0) Then

                Set Component = Components.Item(i)

                'Vapour molar flow of component i
                mi = Stream.VapourPhase.MolarFlows.Values(i) 'or
                'Liquid molar flow of component i
                mi = Stream.LightLiquidPhase.MolarFlows.Values(i) 'or
                'Aqueous molar flow of component i
                mi = Stream.HeavyLiquidPhase.MolarFlows.Values(i)

                'Total molar flow of component i
                mi_T = Stream.MolarFlows.Values(i)

                'Ratios of the components in the vapour phase
                ratioi = mi/mi_T

                'Component's total mole fraction
                moleFraci_T = Stream.MolarFractionsValue(i)

                'Component's phase mole fraction xi
                moleFracPhasei = ratioi * moleFraci_T

                'Calculating the phase specific chemical exergy using std bch
                'from Simulation Basis, xi and the stream molar flow
                bchPhase = bchPhase +
                moleFracPhasei * Component.GetUserProperty("bch") *
                Stream.MolarFlow.GetValue("kgmole/h") 'or

                ② { bchPhase = bchPhase +
                    moleFracPhasei * Component.GetUserProperty("bchl") *
                    Stream.MolarFlow.GetValue("kgmole/h")

            End If
        Next

        ActiveVariableWrapper.Variable.SetValue(bchPhase)

    End If
ErrorHandler:
End Sub

```

Figure C.2: The *user variable* created to calculate the chemical exergy per phase

APPENDIX D

NORMAL OPERATING CONDITIONS

As small solver variations in HYSYS[®] are seen between runs, under identical operating conditions; ten runs of the normal base case simulation were executed to obtain an average of the energy properties. The normal runs are labelled Normal₁-Normal₁₀, and the collection of average energy properties is here on out referred to as the normal operating condition (NOC). Table D.2, Table D.3, and Table D.4 show the physical exergy, chemical exergy, and energy flow data, respectively. The data was used to calculate the averages that make up the NOC depicted in Table D.1. The NOC and Normal₁ are first mentioned in Chapter 4.

Table D.1: Average physical exergy, chemical exergy, and energy flow making up NOC

Stream no	Exergy [kJ/h]		Energy flow $\dot{E}$ [kJ/h]
	B_{ph}	B_{ch}	
1	211991000	6811680000	-329629228
2	153042000	51577400	-1185124908
3	42119200	17608900	22860070
4	22166700	18697700	-348963640
5	853042500	5970380000	-1840857023
7	207502500	5919913100	-3102596564
8	205658500	5915085000	-1445481368
10	288901400	11343900000	-3032884567
11	300436700	11343900000	-2852584935
12	222734000	10664647698	-3978187110
13	220438400	9800699000	-3950984792
14	2178758	864081324	-27203644
15	117531600	9748519204	-4378237273
16	116458300	8837496000	-2607799760
17	246986	1769064000	-78084797
18	533380	5608290	-1719558002
19	23291660	1767499000	-521559765
20	93166650	7069996000	-2086238813
21	108849400	7069996000	-2066399976
22	25253070	1640237000	-479404858
24	83599810	5428814000	-1587406162

Table D.2: The recorded physical exergy data for the 10 runs and resultant average (NOC)

Stream no	Normal ₁	Normal ₂	Normal ₃	Normal ₄	Normal ₅	Normal ₆	Normal ₇	Normal ₈	Normal ₉	Normal ₁₀	NOC
1	211991000	211991000	211991000	211991000	211991000	211991000	211991000	211991000	211991000	211991000	211991000
2	153042000	153042000	153042000	153042000	153042000	153042000	153042000	153042000	153042000	153042000	153042000
3	42119200	42119200	42119200	42119200	42119200	42119200	42119200	42119200	42119200	42119200	42119200
4	22166700	22166700	22166700	22166700	22166700	22166700	22166700	22166700	22166700	22166700	22166700
5	853040000	853045000	853045000	853045000	853045000	853045000	853040000	853040000	853040000	853040000	853042500
7	207502000	207503000	207503000	207503000	207503000	207503000	207502000	207502000	207502000	207502000	207502500
8	205658000	205659000	205659000	205659000	205659000	205659000	205658000	205658000	205658000	205658000	205658500
10	288628000	289156000	289166000	288980000	288880000	288829000	288915000	288731000	288848000	288881000	288901400
11	300122000	300693000	300720000	300505000	300410000	300353000	300435000	300358000	300367000	300404000	300436700
12	222295000	222803000	222826000	222636000	222534000	222493000	222569000	224152000	222504000	222528000	222734000
13	219998000	220508000	220533000	220341000	220241000	220200000	220273000	221846000	220209000	220235000	220438400
14	2180750	2177780	2175880	2178130	2175960	2176180	2178360	2190030	2178210	2176300	2178758
15	117153000	117621000	117644000	117471000	117399000	117350000	117414000	118511000	117356000	117397000	117531600
16	116085000	116549000	116571000	116400000	116328000	116279000	116344000	117414000	116286000	116327000	116458300
17	246183	246103	246176	246119	246076	246139	246062	254831	246157	246018	246986
18	532825	532244	532125	532536	532282	532355	532411	542273	532470	532278	533380
19	23217000	23309800	23314300	23280100	23265600	23255800	23268800	23482700	23257200	23265300	23291660
20	92867900	93239300	93257200	93120300	93062400	93023200	93075100	93930900	93028800	93061400	93166650
21	108638000	109070000	109089000	108931000	108862000	108816000	108879000	108523000	108824000	108862000	108849400
22	25203900	25304300	25308700	25271900	25256100	25245400	25259900	25177300	25247200	25256000	25253070
24	83338600	83865100	83872000	83688400	83587000	83538300	83624500	83337600	83556500	83590100	83599810

Table D.3: The recorded chemical exergy data for the 10 runs and resultant average (NOC)

Stream no	Normal ₁	Normal ₂	Normal ₃	Normal ₄	Normal ₅	Normal ₆	Normal ₇	Normal ₈	Normal ₉	Normal ₁₀	NOC
1	6811680000	6811680000	6811680000	6811680000	6811680000	6811680000	6811680000	6811680000	6811680000	6811680000	6811680000
2	51577400	51577400	51577400	51577400	51577400	51577400	51577400	51577400	51577400	51577400	51577400
3	17608900	17608900	17608900	17608900	17608900	17608900	17608900	17608900	17608900	17608900	17608900
4	18697700	18697700	18697700	18697700	18697700	18697700	18697700	18697700	18697700	18697700	18697700
5	5970380000	5970380000	5970380000	5970380000	5970380000	5970380000	5970380000	5970380000	5970380000	5970380000	5970380000
7	5920444560	5920454550	5920454550	5920454550	5915090000	5920454550	5920444560	5920444560	5920444560	5920444560	5919913100
8	5915080000	5915090000	5915090000	5915090000	5915090000	5915090000	5915080000	5915080000	5915080000	5915080000	5915085000
10	11330900000	11357400000	11353500000	11331400000	11344300000	11340400000	11348800000	11343600000	11343500000	11345200000	11343900000
11	11330900000	11357400000	11353500000	11331400000	11344300000	11340400000	11348800000	11343600000	11343500000	11345200000	11343900000
12	10651174000	10677656000	10675008260	10655449720	10664576000	10660607000	10669114000	10663840000	10663693000	10665359000	10664647698
13	9787070000	9814320000	9811020000	9788350000	9801700000	9797840000	9805610000	9798530000	9800210000	9802340000	9800699000
14	864104000	863336000	863988260	867099720	864202260	862767000	863504000	865310000	863483000	863019000	864081324
15	9734878650	9762140700	9758846860	9736170530	9749529250	9745656560	9753425270	9746350720	9748028460	9750165040	9748519204
16	8823660000	8850980000	8847910000	8825640000	8838960000	8834880000	8842470000	8833960000	8836930000	8839570000	8837496000
17	1769930000	1769100000	1768210000	1768430000	1768050000	1768150000	1769070000	1772300000	1769190000	1768210000	1769064000
18	5607760	5607810	5607920	5608520	5607350	5607650	5607410	5613720	5607580	5607180	5608290
19	1764730000	1770200000	1769580000	1765130000	1767790000	1766980000	1768490000	1766790000	1767390000	1767910000	1767499000
20	7058930000	7080780000	7078330000	7060510000	7071170000	7067910000	7073970000	7067160000	7069540000	7071660000	7069996000
21	7058930000	7080780000	7078330000	7060510000	7071170000	7067910000	7073970000	7067160000	7069540000	7071660000	7069996000
22	1637670000	1642740000	1642170000	1638040000	1640510000	1639750000	1641160000	1639580000	1640130000	1640620000	1640237000
24	5415780000	5442360000	5438370000	5416350000	5429250000	5425340000	5433750000	5428470000	5428400000	5430070000	5428814000

Table D.4: The recorded energy flow data for the 10 runs and resultant average (NOC)

Stream no	Normal ₁	Normal ₂	Normal ₃	Normal ₄	Normal ₅	Normal ₆	Normal ₇	Normal ₈	Normal ₉	Normal ₁₀	NOC
1	-329629228	-329629228	-329629228	-329629228	-329629228	-329629228	-329629228	-329629228	-329629228	-329629228	-329629228
2	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908	-1185124908
3	22860070	22860070	22860070	22860070	22860070	22860070	22860070	22860070	22860070	22860070	22860070
4	-348963640	-348963640	-348963640	-348963640	-348963640	-348963640	-348963640	-348963640	-348963640	-348963640	-348963640
5	-1840863225	-1840850822	-1840850822	-1840850822	-1840850822	-1840850822	-1840863225	-1840863225	-1840863225	-1840863225	-1840857023
7	-3102598631	-3102594497	-3102594497	-3102594497	-3102594497	-3102594497	-3102598631	-3102598631	-3102598631	-3102598631	-3102596564
8	-1445484455	-1445478281	-1445478281	-1445478281	-1445478281	-1445478281	-1445484455	-1445484455	-1445484455	-1445484455	-1445481368
10	-3022827873	-3036801821	-3042616569	-3039294622	-3033182218	-3032891318	-3029560928	-3030888564	-3029783857	-3030997900	-3032884567
11	-2842883716	-2856479289	-2862224260	-2859089360	-2852991157	-2852772983	-2849397057	-2849522727	-2849651038	-2850837767	-2852584935
12	-3968468681	-3982090322	-3987818738	-3984683617	-3978610691	-3978365621	-3975001135	-3975110765	-3975254999	-3976466535	-3978187110
13	-3941248368	-3954889529	-3960640285	-3957485713	-3951428665	-3951184209	-3947797778	-3947845688	-3948055441	-3949272244	-3950984792
14	-27220171	-27198582	-27179011	-27197737	-27186780	-27182286	-27203365	-27274015	-27203515	-27190974	-27203644
15	-4368268686	-4382324593	-4388099033	-4384776432	-4378606884	-4378364277	-4375020386	-4375249189	-4375234375	-4376428879	-4378237273
16	-2597693665	-2611874261	-2617825822	-2614348045	-2608375172	-2608051598	-2604612990	-2604225230	-2604812551	-2606178265	-2607799760
17	-78087931	-78069322	-78051690	-78050720	-78030431	-78039364	-78057670	-78361008	-78070118	-78029714	-78084797
18	-1719711891	-1719569878	-1719412086	-1719574760	-1719380528	-1719459424	-1719564997	-1719940270	-1719554099	-1719412086	-1719558002
19	-519538112	-522374852	-523564542	-522869609	-521675034	-521610320	-520922598	-520844424	-520962510	-521235653	-521559765
20	-2078149961	-2089499409	-2094256924	-2091478436	-2086700138	-2086438790	-2083689149	-2083380184	-2083852527	-2084942612	-2086238813
21	-2058207304	-2069477455	-2074232259	-2071482600	-2066716547	-2066463915	-2063700292	-2064887821	-2063873535	-2064958033	-2066399976
22	-477504833	-480119041	-481222698	-480583618	-479477943	-479420442	-478778640	-479053383	-478818192	-479069796	-479404858
24	-1577344080	-1591320506	-1597138174	-1593817172	-1587707363	-1587419221	-1584083564	-1585410368	-1584300644	-1585520532	-1587406162

APPENDIX E

CALCULATING THE THRESHOLD VALUE

The appendix is used to show how the threshold value (κ) was calculated. κ is used within the threshold function detailed in Chapter 5. As already mentioned, each time the simulation model was run within HYSYS®, under identical operating conditions, small solver variations were observed. These deviations were used to quantify a threshold κ -value to ensure that the simulation variations would not be mistaken for faulty process conditions.

E.1 CALCULATION

In order to assign an appropriate value to κ , the well-known statistical experimental error was used as a base [171]. In order to calculate the error percentage, the parameters summarised in Table E.1 were required.

Table E.1: Parameters and the formulae used to quantify the error percentage of the simulation variations

Parameter		Formula
Symbol	Description	
p	Statistical significance	-
m	Number of samples	-
df	Degrees of freedom	Degree of freedom
μ	Average	$\frac{1}{m} \sum_{i=1}^m B_i$
σ	Standard deviation	$\sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \mu)^2}$
t_{m-1}	t-value	t critical value from table
e	Error	$t_{m-1} \left(\frac{\sigma}{\sqrt{m}} \right)$
$e\%$	Error percentage	Error in terms of percentage

The statistical significance (p) was assumed to be 0.001. The next step was to run the normal operating condition (NOC) simulation model ten times ($m = 10$); each time recording the physical exergy (B_{ph}) and chemical exergy (B_{ch}) of every stream. With the exergy data logged, the average (μ) and standard deviation (σ) of the streams were calculated, respectively. The critical t-value (t_{m-1}) was determined from a statistical lookup table, where $p = 0.001$ and the degrees of freedom is $m - 1 = 9$. From the table, the t-value was found to be 3.25. Using this t-value, the computed standard deviation and number of samples, the error (e) was calculated for every stream. Lastly, the obtained error values were expressed as percentages ($e\%$).

Table E.2 summarises the aforementioned values for every GTL process stream. When evaluating the error percentages, the largest deviation (of 1.15 %) is seen within stream 17. Consequently, 1.2 % was chosen as a suitable threshold value. Seeing as the threshold technique uses normalised data, the percentage was converted to a normalised value of $\kappa = 0.012$.

Table E.2: Calculating the threshold value κ by using the simulation variations

Stream no	p	m	$m - 1$	μ		σ		t_{m-1}	e		$e\%$	
				B_{ph}	B_{ch}	B_{ph}	B_{ch}		B_{ph}	B_{ch}	B_{ph}	B_{ch}
1	0.001	10	9	211991000	6811680000	0	0	3.25	0	0	0.00	0.00
2	0.001	10	9	153042000	51577400	0	0	3.25	0	0	0.00	0.00
3	0.001	10	9	42119200	17608900	0	0	3.25	0	0	0.00	0.00
4	0.001	10	9	22166700	18697700	0	0	3.25	0	0	0.00	0.00
5	0.001	10	9	853042500	5970380000	2635	0	3.25	2708	0	0.00	0.00
7	0.001	10	9	207502500	5919913100	527	1694672	3.25	542	1741682	0.00	0.03
8	0.001	10	9	205658500	5915085000	527	5270	3.25	542	5417	0.00	0.00
10	0.001	10	9	288901400	11343900000	167978	8419686	3.25	172637	8653250	0.06	0.08
11	0.001	10	9	300436700	11343900000	173042	8419686	3.25	177842	8653250	0.06	0.08
12	0.001	10	9	222734000	10664647698	521421	8034996	3.25	535885	8257889	0.24	0.08
13	0.001	10	9	220438400	9800699000	518087	8636052	3.25	532459	8875618	0.24	0.09
14	0.001	10	9	2178758	864081324	4238	1280058	3.25	4355	1315567	0.20	0.15
15	0.001	10	9	117531600	9748519204	371215	8639036	3.25	381512	8878685	0.32	0.09
16	0.001	10	9	116458300	8837496000	363037	8648197	3.25	373107	8888100	0.32	0.10
17	0.001	10	9	246986	1769064000	2757	1289403	3.25	2833	1325171	1.15	0.07
18	0.001	10	9	533380	5608290	3131	1943	3.25	3218	1997	0.60	0.04
19	0.001	10	9	23291660	1767499000	72584	1729646	3.25	74597	1777627	0.32	0.10
20	0.001	10	9	93166650	7069996000	290376	6918035	3.25	298431	7109943	0.32	0.10
21	0.001	10	9	108849400	7069996000	172212	6918035	3.25	176989	7109943	0.16	0.10
24	0.001	10	9	83599810	5428814000	181300	8423102	3.25	186329	8656761	0.22	0.16

APPENDIX F

EIGENVALUES' STANDARD DEVIATION

The appendix shows how the standard deviation of the eigenvalues, used within Chapter 7, were calculated.

F.1 CALCULATION

Approach III.A in Chapter 7 uses the standard deviation of the eigenvalues to assign a qualitative 0 or 1. To determine the standard deviation of the eigenvalues, the following calculations were done:

1. The individual cost matrices of the NOC node signature matrix as compared to all ten normal conditions (Normal_x¹) were obtained using:

$$C_{\text{NOCNormal}_x} = \sqrt{\sum_{a=1}^A \frac{|G_{\text{NOC}_{ia}} - G_{\text{Normal}_{xja}}|}{range_a}}. \quad (\text{F.1})$$

2. Next, the eigenvalues of these cost matrices were obtained by making use of the MATLAB[®] function `D=eig(C, 'vector');`
3. Lastly, the standard deviation (σ) of each one of the eigenvalues were calculated by applying

$$\sigma = \sqrt{\frac{\sum (\lambda_{(v)} - \mu_{\lambda})^2}{m}} \quad (\text{F.2})$$

where $\lambda_{(v)}$ is each of the eigenvalues ($v = 1, \dots, 18$), μ_{λ} the eigenvalue average, and m the sample number. The eigenvalue data that were used is tabulated in Table F.1, and the calculated standard deviation is given in Table F.2.

¹The ten normal conditions labelled Normal₁ – Normal₁₀.

F.2 DATA

Table F.1: Eigenvalues of cost matrices used to determine the standard deviation

Eigenvalues $[\lambda_{(1)}, ..., \lambda_{(18)}]$									
Normal ₁	Normal ₂	Normal ₃	Normal ₄	Normal ₅	Normal ₆	Normal ₇	Normal ₈	Normal ₉	Normal ₁₀
20.9232	20.9402	20.9483	20.9436	20.9349	20.9346	20.9307	20.9346	20.9311	20.9322
-0.1084	-0.1084	-0.1084	-0.1084	-0.1084	-0.1084	-0.1084	-0.1084	-0.1084	-0.1084
-0.2151	-0.2153	-0.2152	-0.2152	-0.2152	-0.2152	-0.2152	-0.2152	-0.2152	-0.2152
-0.2802	-0.2803	-0.2803	-0.2803	-0.2803	-0.2803	-0.2803	-0.2802	-0.2803	-0.2803
-0.4926	-0.4927	-0.4927	-0.4927	-0.4927	-0.4927	-0.4927	-0.4926	-0.4927	-0.4927
-0.6018	-0.6047	-0.6040	-0.6044	-0.6051	-0.6051	-0.6041	-0.6042	-0.6041	-0.6046
-0.7280	-0.7305	-0.7299	-0.7302	-0.7308	-0.7309	-0.7300	-0.7299	-0.7300	-0.7304
-0.8753	-0.8785	-0.8779	-0.8783	-0.8788	-0.8788	-0.8777	-0.8777	-0.8778	-0.8782
-0.9247	-0.9270	-0.9271	-0.9270	-0.9267	-0.9269	-0.9262	-0.9261	-0.9263	-0.9265
-0.9911	-0.9939	-0.9936	-0.9937	-0.9940	-0.9941	-0.9932	-0.9929	-0.9932	-0.9936
-1.0652	-1.0682	-1.0680	-1.0681	-1.0682	-1.0683	-1.0673	-1.0674	-1.0674	-1.0678
-1.2156	-1.2187	-1.2185	-1.2187	-1.2188	-1.2188	-1.2178	-1.2179	-1.2179	-1.2183
-1.2988	-1.3020	-1.3018	-1.3019	-1.3021	-1.3020	-1.3010	-1.3011	-1.3010	-1.3015
-1.4318	-1.4357	-1.4358	-1.4358	-1.4356	-1.4358	-1.4344	-1.4348	-1.4346	-1.4350
-1.5089	-1.5121	-1.5120	-1.5119	-1.5121	-1.5121	-1.5111	-1.5110	-1.5112	-1.5116
-1.8218	-1.8225	-1.8224	-1.8225	-1.8222	-1.8225	-1.8223	-1.8222	-1.8223	-1.8224
-2.5096	-2.5114	-2.5107	-2.5110	-2.5111	-2.5113	-2.5110	-2.5106	-2.5111	-2.5109
-4.8317	-4.8293	-4.8279	-4.8287	-4.8299	-4.8300	-4.8306	-4.8314	-4.8307	-4.8303

Table F.2: Calculated standard deviation for each eigenvalue

σ	3σ
0.006773	0.040637
0.000001	0.000004
0.000037	0.000220
0.000036	0.000216
0.000034	0.000203
0.000872	0.005235
0.000763	0.004576
0.000951	0.005708
0.000676	0.004057
0.000817	0.004903
0.000887	0.005321
0.000911	0.005468
0.000946	0.005677
0.001156	0.006937
0.000935	0.005609
0.000193	0.001158
0.000479	0.002872
0.001102	0.006612

APPENDIX G

IFAC WORLD CONGRESS 2020 ARTICLE

Exergy graph-based fault detection and isolation of a gas-to-liquids process

Sarita Greyling* George van Schoor**
Kenneth Richard Uren* Henri Marais*

* *School of Electrical, Electronic, and Computer Engineering, Faculty of Engineering, North-West University, Potchefstroom 2531, South Africa (e-mail: sarita.svl@gmail.com/*

kenny.uren@nwu.ac.za/henri.marais@nwu.ac.za).

** *Unit for Energy and Technology Systems, Faculty of Engineering, North-West University, Potchefstroom 2531, South Africa (e-mail: george.vanschoor@nwu.ac.za)*

Abstract: With the sheer size of modern process plants, the Fault Detection and Isolation (FDI) field continues to gain popularity. FDI is a sophisticated concept which aims to detect and isolate anomalies that occur within a plant to avoid losses of personnel, damages to the environment, and financial implications. It does so in a way which is more direct, efficient and safer than what human operators are capable of. One approach to FDI is to consider the exergy characterisation of a system. By describing the exergy of the system units and streams, in this case a gas-to-liquids (GTL) process plant, the various process variables are encapsulated under a universal energy-domain parameter. The advantage of this being that it can describe the physical states as well as the chemical characteristics of the process. Previous work which utilised exergy characterisation along with a fixed-threshold approach showed promise. This study however, shows that the approach falls short when presented with 3 % faults. These results motivated the investigation of utilising attributed graphs, which package exergy data into a framework that preserves structural information of the plant. The usefulness of finding similarities (called graph matching) between the graphs constructed of operational conditions and pre-collected fault conditions to detect and isolate faulty conditions, is demonstrated. The technique performs well when considering fault magnitudes bigger than 8 % but deteriorates when applied to smaller, 3 % faults. The poor performance could be ascribed to the graph matching aspect, which is described by a single distance value that discards dimensionality. Future work will therefore look into the graph matching technique specifically, aiming to retain more informative dimensions.

Keywords: Fault detection, Fault isolation, Exergy, Gas-to-liquids, Graph matching, Attributed graphs

1. INTRODUCTION

In most industrial process plants, operators are tasked with the monitoring and management of operations. This means overseeing a large number of units and associated process variables. When anomalies occur within the plant, the operators are expected to detect, diagnose and rectify the situation in the shortest possible time. As technology progresses these industrial processes are enhanced, resulting in even more complex systems. Consequently, the operators' responsibilities could escalate beyond their capabilities. Sometimes the mishandling of events by operators result in costly incidents, not only risking human life and the environment, but also causing detrimental financial situations. Two well-known incidents that illustrate this, is the methyl isocyanate (MIC) leak in Bhopal, India which claimed thousands of lives, according to Kletz (1998). The second incident, as highlighted by Venkatasubramanian et al. (2003), was the Kuwait Petrochemical Mina al Ahmadi oil refinery explosion

which resulted in \$100 million in damages. This is where Fault Detection and Diagnosis (FDD) schemes are of interest. By employing an appropriate FDD approach, the efforts required of operators are reduced, the rectification of anomalies are more efficient and the associated cost and health risks are limited. The advancement of FDD approaches would therefore specifically benefit volatile and expensive processes such as seen in the petrochemical industries (PCIs). FDD approaches are generally categorised as being either model-based or data-driven; the main difference being whether an analytical model is present. Model-based methods utilises analytical or structural models of a process and its behaviours, encapsulating both normal and faulty aspects. Data-driven techniques do not make use of explicit models; but rather derive mathematical relations, based on provided historical process data, between faults and the effects thereof. When surveying literature pertaining to chemical processes, the approaches seem to lean towards being data-driven. The most prominent approaches being either

statistical, as applied by Choi et al. (2005), Xie et al. (2013), Ghosh et al. (2014), Fezai et al. (2018), and Dong and Qin (2018); or machine learning as demonstrated by Watanabe and Hirota (1991), and Sorsa et al. (1991), to name but a few. Venkatasubramanian et al. (2003) states that it would not be impossible to develop analytical models of petrochemical (PC) processes but that it would be exceptionally challenging. Recent research, such as done by Chiang and Braatz (2003), Maki et al. (2004) and Chiang et al. (2015), show hybrid approaches taking the forefront. The hybridisation, which is usually a combination of model-based and data-driven techniques, endeavours to leverage the advantages afforded by the different approaches whilst minimising the drawbacks. Of noticeable interest is the approach proposed by Chiang and Braatz (2003) which aimed at combining causal maps and Partial Least Square (PLS) methods in order to include process connectivity information. Much in the same direction of thinking, Marais et al. (2019), proposed a hybrid approach which makes use of energy properties of the system. Not only is the energy description a unifying parameter across different domains, but it is also a way of abstracting data. The energy properties are then packaged in such a manner that the physical structural information of the system is retained. In the work of Greyling et al. (2019) the same approach was applied to a gas-to-liquids (GTL) process, incorporating exergy rather than energy. By monitoring the exergy, additional information is encapsulated, specifically regarding the chemical characteristics of the system. The results recorded in the work done by Greyling et al. (2019) showed promise, however the question that arose was whether a fixed-threshold approach would still work if the system faults' magnitudes were smaller.

This paper is divided into two parts. The first part focusses on evaluating the threshold approach performance when presented with 3 % faults. The results proved to be less than stellar, which meant some alterations were required. It must however be emphasised that the exergy and structural information concepts still hold promise; the issue seemed to be the fixed-threshold applied. Therefore, keeping with the exergy characterisation and wanting to preserve the structural information, Ould-Bouamama et al. (2014) suggest that a graphical method would allow for both. Such graphical approaches would also provide different mathematical schemes of detecting and isolating considered faults. Most of the graphical approaches reviewed by Ould-Bouamama et al. (2014) make use of graphs to describe system properties and relevant causal relations. For this study the most suitable graphical approach was chosen to be attributed graphs along with graph matching, a popular technique that quantifies the dissimilarities of compared graphs. The second part of the paper therefore demonstrates the usefulness of comparing operational graphs to faulty graphs (stored in a database) in order to detect and isolate faulty conditions.

The paper starts off with briefly detailing the GTL model and exergy characterisation thereof. Section 3 goes on to describe the considered faults' detail and their locations. The threshold approach as applied to 3 % faults and the results obtained is summarised in Section 4. Section 5 gives a quick overview of attributed graphs and graph matching

and goes on to detail the methodology as these are applied to the GTL process. The fault detection and isolation results obtained are given in Section 6 with the paper being concluded in Section 7.

2. THE GAS-TO-LIQUIDS PROCESS

A gas-to-liquids (GTL) process is used to transform gaseous feedstock, such as natural gas, into hydrocarbon liquids. A GTL process usually comprises of three principal sections as shown in Fig. 1. In the first section the natural gas is reformed to obtain synthesis gas, also referred to as syngas. The syngas is made up of a certain ratio of hydrogen (H_2) and carbon monoxide (CO), depending on the intended end-products. The produced syngas is then introduced to a Fischer-Tropsch reactor which converts the syngas into an array of hydrocarbons (syncrude). The upgrading section is used to rework the syncrude to hydrocarbon products of particular chain lengths. Since the upgrading section is quite complex only the first two sections of the GTL process, shown boxed in Fig. 1, are considered in this study. Interested readers are referred to the works of Rafiee and Hillestad (2010), Panahi et al. (2011), De Klerk (2011), and Knutsen (2013) for comprehensive information on the GTL process and the modelling thereof.

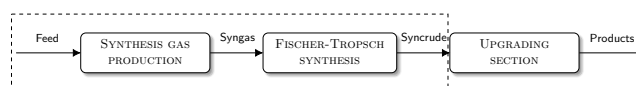


Fig. 1. A process diagram of a gas-to-liquids (GTL) process

2.1 Simulation model

In order to have a representative system to work with, a **steady-state** simulation model was constructed in the commercial process simulator, Aspen HYSYS[®]. No process variations were intentionally included in this study. The exact particulars on how the model was developed, the operating points and validation of the model are comprehensively documented in Greyling et al. (2019). However, some of the most important modelling aspects are highlighted as:

- (1) Autothermal reformer (ATR)
 - (a) No pre-reformer was included as there was no recycling to the ATR.
 - (b) The feedstocks used were pure methane (CH_4), steam ($H_2O(g)$), oxygen (O_2) and carbon dioxide (CO_2).
- (2) Fischer-Tropsch reactor (FTR)
 - (a) A plug flow reactor in conjunction with a separator was used to represent a multi-tubular fixed bed (MTFB) reactor.
 - (b) The syngas that was fed into the FTR was at a temperature of 210 °C.
- (3) Recycle
 - (a) 76.8 % of the unreacted syngas in stream 16 was recycled back to the FTR.
 - (b) The remaining 23.2 % was purged (stream 22).

Fig. 2 shows the Aspen HYSYS[®] process flow diagram of the developed GTL process. Note that the validation of

the simulation model comprised of comparing the attained product distribution to the theoretical distributions seen in literature.

2.2 Exergy characterisation

According to Dincer and Rosen (2013), exergy is defined as being a quantitative measure of an energy quantity's usefulness to perform work. Unlike energy which is based only on the first law of thermodynamics, exergy also takes into account the second law of thermodynamics. The second law states that entropy cannot decrease in any real process, therefore the ability to deliver valuable work is eventually lost. In other words exergy is not conserved and some exergy losses would occur which could be quantified by using the process' exergy balance (Maganelli et al. (2018)). Therefore the most prominent advantage of using exergy is that it enables a manner of quantifying the quality of an energy stream or (more importantly) the efficiency of various elements. Consequently, any deviation of the known efficiencies could be indicative of an anomaly within the system. To characterise the GTL system the intrinsic exergy of each stream was calculated. It is important to note that exergy is always evaluated relative to a reference environment (RE). This means that the RE's intensive properties will determine the exergy. For physical exergy, these include temperature and pressure and are typically $T_0 = 25^\circ\text{C}$ and $P_0 = 101.325\text{ kPa}$. However, the chemical exergy is based on an environment consisting of certain reference elements. Various methods exist for selecting and calculating the standard chemical exergy (b_{ch}^0) of these reference elements with the RE proposed by Szargut (2007) being the most distinguished one. In order to automatically calculate the physical and chemical exergy within Aspen HYSYS[®], *user variables* were developed. A *user variables* is a section of program code that the user creates, which can access various elements of the Aspen HYSYS[®] model. To calculate the physical exergy per mole

$$b_{ph} = (h - h_0) - T_0(s - s_0), \quad (1)$$

is used where h and s are the current enthalpy and entropy, respectively, and h_0 and s_0 the enthalpy and entropy at RE state. The total physical exergy (B_{ph}) is obtained by multiplying the stream's molar flow rate with the computed intrinsic physical exergy. The chemical exergy is calculated by making use of

$$b_{ch} = \sum x_{(i)} b_{ch(i)}^0, \quad (2)$$

with $x_{(i)}$ being the mole fraction and $b_{ch(i)}^0$ the standard molar chemical exergy of substance i . The utilised $b_{ch(i)}^0$ values were defined by Szargut's (2007). In order to utilise (2), the standard molar chemical exergies of all the relevant substances were firstly made available to the *simulation basis* by creating a *user property* tabulating the corresponding values. Not all substances' standard chemical exergies were readily available but were pre-calculated (Greyling et al. (2019)). Seeing as the GTL process would inevitably have multi-phase streams and since some substances have different standard chemical exergies based on their phase, Equation (2) was modified to account for this. Thus the total chemical exergy is the sum of the vapour, liquid and aqueous phases mathematically expresses as

$$b_{ch} = \sum x_{(i)v} b_{ch(i)v}^0 + \sum x_{(i)\ell} b_{ch(i)\ell}^0 + \sum x_{(i)a} b_{ch(i)a}^0. \quad (3)$$

Whenever a certain phase was not present, the phase exergy was assumed to be zero. The complete details on how the physical and chemical exergy *user variables* were developed is also discussed in Greyling et al. (2019). The assumption was made that the physical exergy (B_{ph}) and chemical exergy (B_{ch}) of all the streams and units were available to utilise. The next iteration of the research will look into using fewer data points that carry more importance.

3. SYSTEM FAULT SPECIFICATIONS

Before the considered faults are introduced, a recap of the terminology is crucial. According to Shah (2011), a *failure* is a permanent disruption of the system's operations. A *disturbance* is an unknown input that negatively affects the system's performance where a *fault* is any unintentional deviation of a parameter from its normal behaviour. A fault can be further classified based on its physical location or the effects on the system operation. Faults classified based on their location are system faults, actuator faults, and sensor faults. The different fault effects can be seen as additive, multiplicative, abrupt, incipient, intermittent, permanent, or transient. In order to evaluate the fault detection capabilities of the proposed approaches, eleven relevant fault conditions were identified. These eleven faults include system faults and actuator faults. For this study however, the proposed system need not distinguish between the two fault-categories. The reasoning behind how the faults were chosen, was established in Greyling et al. (2019). The details of faults and their location are summarised in Table 1. The locations of the faults are also visually depicted using danger triangles, as shown in Fig. 2. In order to keep track of the magnitude variation of the faults, fault IDs were assigned to each. The general form of the fault ID is given as F_{pq_r} , where p denotes the relevant GTL section, q the type of fault within the section and r the magnitude deviation considered; the magnitude variations being 3 %, 8 %, 9 %, 10 %, 11 %, 12 %, and 25 %. The most important datasets to take note of are:

- F_{pq_1} are the eleven faults that deviated with a magnitude of 3 %.
- F_{pq_4} are the eleven faults that deviated with a magnitude of 10 %.
- F_{pq_R} is a random selection of various magnitude deviations, *excluding* 3 and 10 % magnitudes, of each one of the eleven faults.

The specified datasets are shown highlighted in Table 1 for easy identification. The GTL model was modified individually and simulated for every fault tabulated, each time recording all the streams' physical exergy (B_{ph}) and chemical exergy (B_{ch}) to use as the exergy characterisation information.

4. A THRESHOLD APPROACH

It is important to note that the threshold approach developed, evaluated and the results documented in Greyling et al. (2019) were considering the 10 %

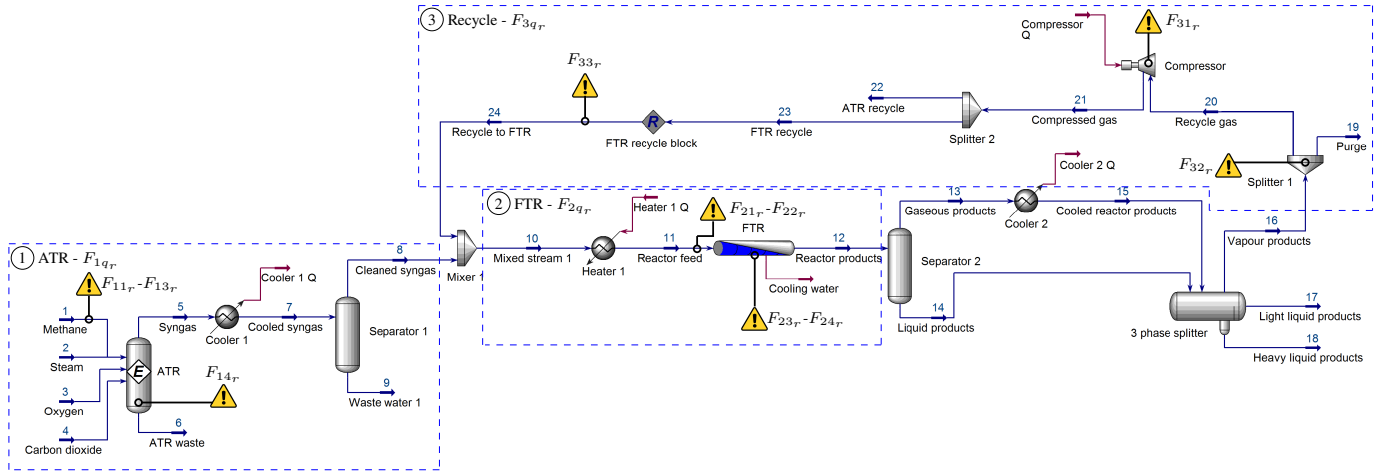


Fig. 2. The GTL process as developed in Aspen HYSYS[®] with fault locations indicated

Table 1. The faults' details and denotation

Fault ID [†]	Description	r							Location	
		1	2	3	4	5	6	7		
ATR section										
F_{1q_r}										
F_{11_r}	Molar flow	+	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Methane stream
F_{12_r}	Molar flow	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Methane stream
F_{13_r}	Pressure	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Methane stream
F_{14_r}	Pressure	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	ATR
FTR section										
F_{21_r}	Temperature	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Reactor feed stream
F_{22_r}	Leakage	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Reactor feed stream
F_{23_r}	Pressure	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	FTR
F_{24_r}	Temperature	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	FTR
Recycle section										
F_{31_r}	Pressure	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Compressor
F_{32_r}	Lower split ratio	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Splitter 1
F_{33_r}	Leakage	-	3 %	8 %	9 %	10 %	11 %	12 %	25 %	Recycle to FTR

Datasets F_{pq1} , F_{pq4} and F_{pqR}

[†] $F_{pq,r}$ - p represents the section, q the fault number and r the magnitude deviation

(F_{pq4}) deviations only. The approach performed very well, successfully detecting all eleven faults and the resultant isolability being 100 %. *Detection* being able to correctly indicate that a fault was present and *isolability* specifically referring to whether the faults were uniquely distinguishable from one another. Subsequently the question arose as to how well the approach would perform when small fault magnitudes are evaluated. To determine this, the following methodology was applied to the 3 % dataset (F_{pq1}):

- (1) Firstly the collected exergy data, per stream, was normalised with respect to the normal condition.
- (2) Next a simple threshold function was applied to the normalised values in order to obtain a qualitative value for each entry. The threshold function used is described by:

$$y = \begin{cases} -1 & \text{if } z < (1 - \kappa) \\ 1 & \text{if } z > (1 + \kappa) \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

In (4), z represents the normalised exergy value being considered and y the magnitude of the resultant fault element. In order to assign an appropriate value to κ , the solver deviations seen within the Aspen HYSYS[®] environment were utilised. As

the simulation model was rerun - under identical conditions - small solver variations were noticed. To ensure that these simulation variations were distinguishable from the faults, the variances were quantified. This was achieved by calculating the statistical experimental error between 10 simulation runs. The threshold κ -value was found to be 0.00635; the precise calculation hereof shown in Greyling et al. (2019).

- (3) After applying the threshold function to the normalised data, a 20×2 qualitative matrix is obtained with the form

$$F_{pq,r} = \begin{bmatrix} y_{B_{ph}(stream1)} & y_{B_{ch}(stream1)} \\ \vdots & \vdots \\ y_{B_{ph}(stream24)} & y_{B_{ch}(stream24)} \end{bmatrix}. \quad (5)$$

- (4) When evaluating the detection and isolation performance
 - (a) any non-zero matrix would indicate a fault condition
 - (b) any identical matrices, for different fault conditions, would signify unisolability

Table 2 shows the qualitative matrices obtained for dataset F_{pq1} after applying the threshold function. When evaluating the matrices, a shortcoming in terms of detection is evident. Seeing that the qualitative matrix of fault $F_{23,1}$ is zero, the fault condition was not successfully detected. To calculate the detection rates of a proposed approach, a confusion matrix is drawn up. The idea behind a confusion matrix is to determine the number of instances where the decision of the approach:

- resulted in a *true negative* (TN) - the approach detected a fault-free condition and the true condition was indeed fault-free (value assigned to a)
- resulted in a *false negative* (FN) - the approach detected a fault-free condition and the true condition was faulty (value assigned to b)
- resulted in a *false positive* (FP) - the approach detected a fault condition and the true condition was actually fault-free (value assigned to c)

Table 2. Threshold results for dataset F_{pq_1}

Stream	F_{1q_1}								F_{2q_1}								F_{3q_1}							
	F_{11_1}		F_{12_1}		F_{13_1}		F_{14_1}		F_{21_1}		F_{22_1}		F_{23_1}		F_{24_1}		F_{31_1}		F_{32_1}		F_{33_1}			
	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}	B_{ph}	B_{ch}		
1	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
5	0	1	1	-1	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
7	1	1	-1	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
8	1	1	-1	-1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
10	1	1	-1	-1	1	0	0	0	0	-1	-1	0	0	0	0	-1	0	-1	-1	-1	-1	-1		
11	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	0	0	0	0	0	-1	-1	-1	-1	-1		
12	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	0	0	1	0	0	-1	-1	-1	-1	-1		
13	1	1	-1	-1	1	0	0	0	-1	0	-1	-1	0	0	1	1	0	0	-1	-1	-1	-1		
14	1	-1	-1	1	1	-1	1	-1	-1	1	1	0	0	1	-1	0	0	1	1	1	1	1		
15	1	1	-1	-1	0	0	0	0	0	-1	-1	0	0	0	1	0	0	-1	-1	-1	-1	-1		
16	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	0	-1	-1	-1	-1	-1		
17	-1	-1	0	0	1	1	0	0	-1	0	1	-1	0	0	0	0	0	0	0	0	1	1		
18	1	1	-1	-1	1	1	0	0	-1	0	1	1	0	0	1	0	0	0	1	0	1	0		
19	1	1	-1	-1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	-1	-1	-1		
20	1	1	-1	-1	0	0	0	0	0	0	1	-1	0	0	0	0	0	-1	-1	-1	-1	-1		
21	1	1	-1	-1	0	0	0	0	0	0	-1	-1	0	0	0	0	-1	0	-1	-1	-1	-1		
24	1	1	-1	-1	0	0	0	0	0	0	1	1	0	0	0	0	-1	0	-1	-1	-1	-1		

- resulted in a *true positive* (TP) - the approach detected a fault condition and the true condition was faulty (value assigned to d)

Subsequently, the detection rates r_{FP} , r_{FN} , r_{TP} , and accuracy are calculated by making use of these assigned values. Ideally, the false positive rate (r_{FP}) and false negative rate (r_{FN}) should be 0 % and the true positive rate (r_{TP}) and accuracy 100 %. A confusion matrix is completed for the threshold approach and is shown in Fig. 3. Evaluating these obtained rates, it is clear that the threshold approach did not perform flawlessly, therefore motivating the development of a different approach.

CONFUSION MATRIX				DETECTION RATES		
		TRUE CONDITION		Rate	Formula	%
		Fault-free	Fault			
DETECTION CONDITION	Fault-free	True negative a	False negative b	r_{FP}	$= \frac{b}{(a+c)} \times 100$	0
	Fault	False positive c	True positive d	r_{FN}	$= \frac{c}{(b+d)} \times 100$	9.09
				r_{TP}	$= \frac{d}{(b+d)} \times 100$	90.91
				Accuracy	$= \frac{(a+d)}{(a+b+c+d)} \times 100$	90.91

Fig. 3. Confusion matrix and detection rates of threshold approach applied to dataset F_{pq_1}

5. GRAPH THEORETICAL APPROACH

5.1 Background

Graph theory has been in use since the 1730's and became very popular in the 1930's. It is mathematical in nature and the concepts thereof have diverse capabilities. A broad spectrum of applications are seen throughout literature, including pattern recognition, transportation and even economics. A graph essentially consists of an ordered pair $G = (V, E)$, where V is the set of vertices (also called nodes) and E the set of edges (sometimes referred to as links or arcs). Usually vertices represent certain properties of a system, whereas the edges are used to describe the incidence relation of the vertices to themselves or other vertices within the graph, as stated by Bondy

and Murty (1976). Furthermore, the graph vertices and edges can contain information. If the information is simply a name or number, the graph is called a *labelled graph*. Should additional information in the form of attributes be available, the graph is called an *attributed graph*. The edges can also be either directional or have no direction related to it. From the definition it is evident then why graph theory can be utilised in so many fields, notwithstanding FDD.

To show how one would go about constructing an attributed graph of the GTL process, the ATR unit will be used as an illustrative example. Fig. 4 (a) shows the ATR unit with its feed streams and syngas product stream. Firstly, each feed stream is represented by a node (nodes 1–4). These nodes are then connected to the ATR unit node (node 5) via directed edges, just as the process flow diagram depicts. The attributes of the nodes and edges are described by the energy characteristics calculated of the process. The completed attributed graph of the ATR unit is shown in Fig. 4 (b). An invaluable aspect of graph theory, called graph matching, is the determination of how similar one graph is to another. As summarised in Wilson and Martinez (1997), many different matching methods exist and the technique applied greatly depends on the type of graphs, their sizes, and the relevant information (symbolic, numeric,

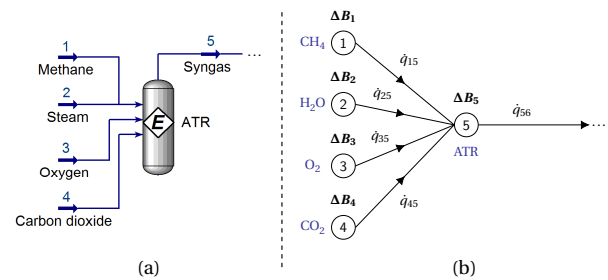


Fig. 4. (a) The ATR process unit (b) The constructed graph showing the nodes, edges and energy attributes of the ATR

etc.) being considered. For this study the Heterogeneous Euclidean Overlapping Metric (HEOM) proposed by Wilson and Martinez (1997), and reiterated by Jouli et al. (2009), was used. The technique works for both numerical and symbolic attributes, although the attributes in this case are numerical only. By utilising the HEOM instead of just calculating the Euclidean distance the following aspects are addressed:

- Should symbolic attributes be included in the future, the HEOM approach will be able to adequately handle the additional information.
- The Euclidean distance function does not include any normalisation, therefore, according to Wilson and Martinez (1997), attributes with large ranges would diminish smaller attributes' inputs.

5.2 Methodology

This section details the methodology of applying graph matching as a means to fault detection and isolation. To ensure a repeatable procedure, the following steps were determined and applied:

- (1) An attributed graph, as depicted in Fig. 5, was constructed based on the GTL process flow diagram, where the:
 - (a) *nodes* represent the process units
 - (b) *edges* convey the flow direction and connection of the units
- (2) The attributes of the:
 - (a) *nodes* are the changes in physical and chemical exergy (ΔB) over each process unit
 - (b) *edges* are the energy flows ($\dot{q}_{\iota\gamma}$) between connected process units ι and γ
- (3) Utilising the graph, a node signature matrix $\mathbf{G}$ is constructed in the form given in (6); describing the change in physical exergy (ΔB_{ph}), chemical exergy (ΔB_{ch}) and the energy flow ($\dot{q}_{\iota\gamma}$) of each node. Should there be no energy flowing between two nodes, a 0 is added to that entry. Matrices were developed for each fault in datasets F_{pq_4} , F_{pq_R} , and F_{pq_1} .

$$\mathbf{G} = \begin{bmatrix} \Delta B_{ph_1} & \Delta B_{ch_1} & \dot{q}_{11} & \dots & \dot{q}_{118} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta B_{ph_{18}} & \Delta B_{ch_{18}} & \dot{q}_{181} & \dots & \dot{q}_{1818} \end{bmatrix} \quad (6)$$

- (4) Next a database was developed containing the graphs ($\mathbf{G}_d$) of every fault of dataset F_{pq_4} . No graph information ($\mathbf{G}_o$) pertaining to the operational faults

to be evaluated (F_{pq_R} and F_{pq_1}) are included in the database.

- (5) A cost matrix $\mathbf{C}_{od}$ is used to determine how dissimilar two graphs, $\mathbf{G}_o$ and $\mathbf{G}_d$, are when compared to one another. To calculate this

$$\mathbf{C}_{od} = \sqrt{\sum_{\alpha=1}^A \frac{|\mathbf{G}_{o_{i\alpha}} - \mathbf{G}_{d_{j\alpha}}|^2}{range_{\alpha}}}, \quad (7)$$

is used, giving an $(i \times j)$ matrix. A is the number of columns of the graphs, j the number of rows in graph $\mathbf{G}_d$ and i the number of rows in graph $\mathbf{G}_o$. To normalise the data, the $range_{\alpha}$ of each column of graph $\mathbf{G}_o$ is obtained and calculated by:

$$range_{\alpha} = max_{\alpha} - min_{\alpha}, \quad (8)$$

where max_{α} is the largest numerical value and min_{α} the smallest in column α .

- (6) In order to determine a single distance (D_C) parameter between the two considered graphs,

$$D_C = \frac{\sum_{k=1}^i \mathbf{C}_{kk}}{i}, \quad (9)$$

is utilised. The calculation basically comes down to summing the cost matrix's diagonal entries and dividing it by the number of rows, i , in the cost matrix.

- (7) The smaller the D_C -value the smaller the dissimilarities are between the compared graphs.

The detection and isolation will therefore work on the premise that given a known fault type of unknown magnitude, the above described method should match the operational condition to the corresponding fault - or more specifically - the graph within the database, by means of the smallest D_C -value.

6. RESULTS

The above-mentioned methodology was firstly applied to fault dataset F_{pq_R} . The D_C -values of each one of the eleven faults as compared to the database stored graphs were recorded and summarised in Table 3 (a). The smallest value, shown in bold, indicates the likeliest match. When evaluating the D_C -values it is seen that the proposed graph matching approach correctly matched all considered operational faults in dataset F_{pq_R} to their corresponding database faults. A confusion matrix was completed and is shown in Fig. 6 (a). The approach performed quite well as there were no false negatives (FN) or false positives

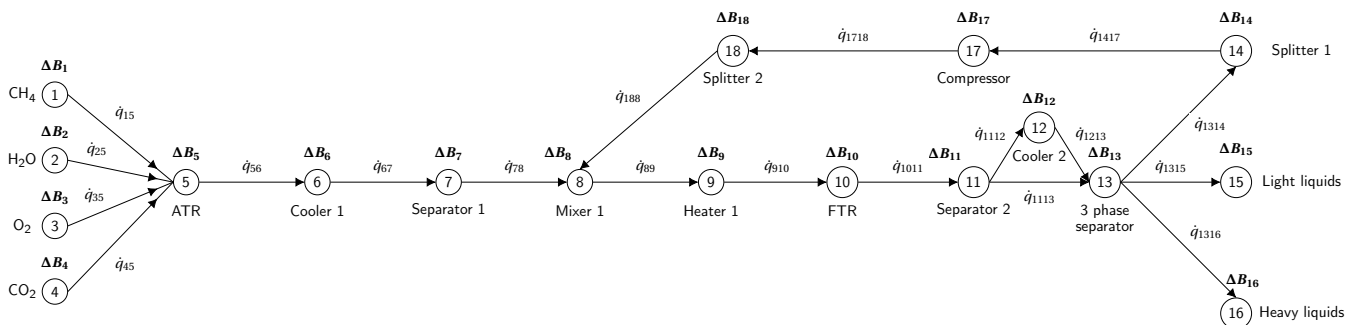


Fig. 5. The graph of the GTL process showing all of the nodes, edges and energy attributes

(FP), with both the true positives (TP) and accuracy being 100 %. However, when the approach was applied to dataset F_{pq_1} the performance drastically deteriorated. The D_C -values, shown in Table 3 (b), show poor matchings of the smaller magnitude faults. The confusion matrix for dataset F_{pq_1} is depicted in Fig. 6 (b). The fact that some faults graphs (F_{14_1} , F_{21_1} , F_{23_1} , F_{24_1} , and F_{31_1}) were matched to the normal graph seems to emphasise an issue

regarding the sensitivity of the proposed approach. The false negative rate (r_{FN}) of 45.5 % and accuracy of 54.5 %, clearly indicate the poor performance.

7. CONCLUSION

As the fixed-threshold approach, proposed in the work of Greyling et al. (2019), failed to detect all the considered 3 % faults, a different detection and isolation method

Table 3. Detectability and isolability of fault dataset (a) F_{pq_R} and (b) F_{pq_1}

(a)

FAULT ID				DATABASE STORED FAULTS											
				Normal	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}
F_{pq_R}	Normal ₁	✓	✓	0.00126	0.05044	0.06362	0.00362	0.00355	0.00260	0.09191	0.00169	0.01641	0.00142	0.05366	0.05086
	F_{11_6}	✓	✓	0.05018	0.00556	0.10589	0.04919	0.04929	0.04998	0.12818	0.05010	0.04997	0.05070	0.09310	0.08899
	F_{12_3}	✓	✓	0.05929	0.10922	0.01049	0.06119	0.06117	0.05901	0.07902	0.05885	0.06326	0.05881	0.05638	0.05315
	F_{13_7}	✓	✓	0.00726	0.04812	0.07038	0.00458	0.00458	0.00890	0.09882	0.00822	0.02095	0.00834	0.06064	0.05793
	F_{14_2}	✓	✓	0.00206	0.04839	0.06624	0.00059	0.00055	0.00405	0.09447	0.00310	0.01725	0.00332	0.05631	0.05355
	F_{21_3}	✓	✓	0.00211	0.04906	0.06420	0.00434	0.00436	0.00017	0.09232	0.00175	0.01659	0.00309	0.05434	0.05148
	F_{22_5}	✓	✓	0.11739	0.16831	0.08116	0.11989	0.11980	0.11688	0.01518	0.11671	0.11340	0.11684	0.06004	0.06787
	F_{23_6}	✓	✓	0.00146	0.04898	0.06414	0.00382	0.00384	0.00206	0.09228	0.00037	0.01571	0.00259	0.05416	0.05132
	F_{24_2}	✓	✓	0.00567	0.05032	0.06647	0.00668	0.00674	0.00767	0.09370	0.00667	0.00456	0.00661	0.05512	0.05257
	F_{31_7}	✓	✓	0.00311	0.04901	0.06532	0.00547	0.00551	0.00511	0.09350	0.00407	0.01204	0.00238	0.05524	0.05250
	F_{32_2}	✓	✓	0.04880	0.09490	0.05312	0.05110	0.05103	0.04844	0.05979	0.04810	0.04990	0.04827	0.01427	0.02527
	F_{33_3}	✓	✓	0.04776	0.09958	0.05292	0.04955	0.04939	0.04811	0.06359	0.04774	0.04595	0.04715	0.02314	0.02178

Yes = ✓ No = ✗

(b)

FAULT ID				DATABASE STORED FAULTS											
				Normal	F_{11_4}	F_{12_4}	F_{13_4}	F_{14_4}	F_{21_4}	F_{22_4}	F_{23_4}	F_{24_4}	F_{31_4}	F_{32_4}	F_{33_4}
F_{pq_1}	F_{11_1}	✓	✗	0.01370	0.03447	0.07613	0.01284	0.01296	0.01385	0.10273	0.01386	0.02045	0.01438	0.06511	0.06207
	F_{12_1}	✓	✗	0.01406	0.06389	0.05221	0.01589	0.01589	0.01434	0.09127	0.01397	0.01833	0.01375	0.05321	0.05049
	F_{13_1}	✓	✓	0.00778	0.04812	0.07087	0.00508	0.00510	0.00939	0.09930	0.00872	0.01341	0.00886	0.06109	0.05837
	F_{14_1}	✗	✗	0.00074	0.04884	0.06515	0.00188	0.00186	0.00287	0.09338	0.00182	0.00977	0.00209	0.05523	0.05245
	F_{21_1}	✗	✗	0.00079	0.04932	0.06427	0.00316	0.00315	0.00150	0.09248	0.00107	0.00976	0.00185	0.05437	0.05156
	F_{22_1}	✓	✗	0.16198	0.18545	0.17694	0.16355	0.16353	0.16160	0.17489	0.16209	0.16144	0.16190	0.16488	0.16488
	F_{23_1}	✗	✗	0.00012	0.04928	0.06446	0.00268	0.00265	0.00223	0.09270	0.00111	0.00941	0.00143	0.05454	0.05177
	F_{24_1}	✗	✗	0.00206	0.04886	0.06506	0.00368	0.00378	0.00373	0.09293	0.00270	0.00842	0.00307	0.05450	0.05179
	F_{31_1}	✗	✗	0.00071	0.04961	0.06416	0.00316	0.00313	0.00234	0.09241	0.00131	0.00938	0.00096	0.05419	0.05139
	F_{32_1}	✓	✗	0.01701	0.06643	0.05301	0.01891	0.01877	0.01752	0.08161	0.01703	0.01907	0.01653	0.04127	0.04062
	F_{33_1}	✓	✗	0.01620	0.06601	0.05405	0.01802	0.01787	0.01673	0.08251	0.01629	0.01728	0.01570	0.04341	0.03985

Yes = ✓ No = ✗

CONFUSION MATRIX				DETECTION RATES		
		TRUE CONDITION		Rate	Formula	%
		Fault-free	Fault			
DETECTION CONDITION	Fault-free	True negative a	False negative b	r_{FP}	$= \frac{c}{(a+c)} \times 100$	0
	Fault	False positive c	True positive d	r_{FN} r_{TP} Accuracy	$= \frac{b}{(b+d)} \times 100$ $= \frac{d}{(b+d)} \times 100$ $= \frac{(a+d)}{(a+b+c+d)} \times 100$	0 100 100

(a)

CONFUSION MATRIX				DETECTION RATES		
		TRUE CONDITION		Rate	Formula	%
		Fault-free	Fault			
DETECTION CONDITION	Fault-free	True negative a	False negative b	r_{FP}	$= \frac{c}{(a+c)} \times 100$	0
	Fault	False positive c	True positive d	r_{FN} r_{TP} Accuracy	$= \frac{b}{(b+d)} \times 100$ $= \frac{d}{(b+d)} \times 100$ $= \frac{(a+d)}{(a+b+c+d)} \times 100$	45.5 54.5 54.5

(b)

Fig. 6. Confusion matrix and detection rates of the graph approach applied to dataset (a) F_{pq_R} and (b) F_{pq_1}

was required. By developing an attributed graph of the GTL process, information regarding the physical structure as well as stream composition and physical properties (described by exergy), are encapsulated. Graph theory then provides an array of methods in which to detect and isolate faults. The proposed approach presumed that the graphs of the 10 % faults were available as a database. The graphs of unknown operational faults are then compared to the database faults and their dissimilarities quantified by means of a distance parameter (D_C). The most likely fault being identified by the smallest D_C -value obtained. When evaluating the detection and isolation performance, all faults in dataset F_{pqR} were successfully detected and isolated. Unfortunately, the performance declines when presented with the smaller faults as in dataset F_{pq1} . This would suggest a similar issue with sensitivity such as the threshold approach displayed. One reason for this could be that the chosen D_C metric discards useful information contained within the cost matrix, seeing as the 18×18 matrix is reduced to a single distance parameter. Hence, the positive performance of the F_{pqR} dataset warrants further investigation into the benefits more degrees of freedom would bring about. Additionally, a next phase of the research would assess the proposed approach's performance when completely unanticipated faults are considered.

ACKNOWLEDGEMENTS

This work is based on the research supported wholly/in part by the National Research Foundation of South Africa (Grant Number 127483). This work is based on the research supported by Sasol (Pty) Ltd. Opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to Sasol.

REFERENCES

- Bondy, J. and Murty, U. (1976). *Graph Theory with Applications*. Elsevier Science Publishing Co., Inc.
- Chiang, L.H. and Braatz, R.D. (2003). Process monitoring using causal map and multivariate statistics: fault detection and identification. *Chemometrics and intelligent laboratory systems*, 65(2), 159–178.
- Chiang, L.H., Jiang, B., Zhu, X., Huang, D., and Braatz, R.D. (2015). Diagnosis of multiple and unknown faults using the causal map and multivariate statistics. *Journal of Process Control*, 28, 27–39.
- Choi, S.W., Lee, C., Lee, J.M., Park, J.H., and Lee, I.B. (2005). Fault detection and identification of nonlinear processes based on kernel PCA. *Chemometrics and intelligent laboratory systems*, 75(1), 55–67.
- De Klerk, A. (2011). *Fischer-Tropsch Refining*. John Wiley & Sons, first edition.
- Dong, Y. and Qin, S.J. (2018). A novel dynamic PCA algorithm for dynamic data modeling and process monitoring. *Journal of Process Control*, 67, 1–11.
- Fezai, R., Mansouri, M., Taouali, O., Harkat, M.F., and Bouguila, N. (2018). Online reduced kernel principal component analysis for process monitoring. *Journal of Process Control*, 61, 1–11.
- Ghosh, K., Ramteke, M., and Srinivasan, R. (2014). Optimal variable selection for effective statistical process monitoring. *Computers & Chemical Engineering*, 60, 260–276.
- Greyling, S., Marais, H., van Schoor, G., and Uren, K.R. (2019). Application of exergy-based fault detection in a gas-to-liquids process plant. *Entropy*, 21(6), 565.
- Jouili, S., Mili, I., and Tabbone, S. (2009). Attributed graph matching using local descriptions. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, 89–99. Springer.
- Kletz, T. (1998). *What Went Wrong? Case Histories of Process Plant Disasters*. Elsevier Science, 4th edition.
- Knutsen, K.T. (2013). *Modelling and optimization of a Gas-to-Liquid plant*. Master's thesis, Institutt for kjemisk prosesssteknologi.
- Magnanelli, E., Berglihn, O.T., and Kjelstrup, S. (2018). Exergy-based performance indicators for industrial practice. *International Journal of Energy Research*, 42(13), 3989–4007.
- Maki, M., Jiang, J., and Hagino, K. (2004). A stability guaranteed active fault-tolerant control system against actuator failures. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 14(12), 1061–1077.
- Marais, H., van Schoor, G., and Uren, K.R. (2019). The merits of exergy-based fault detection in petrochemical processes. *Journal of Process Control*, 74, 110–119.
- Ould-Bouamama, B., Biswas, G., Loureiro, R., and Merzouki, R. (2014). Graphical methods for diagnosis of dynamic systems: Review. *Annual Reviews in Control*, 38(2), 199–219.
- Panahi, M., Rafiee, A., Skogestad, S., and Hillestad, M. (2011). A natural gas to liquids process model for optimal operation. *Industrial & Engineering Chemistry Research*, 51(1), 425–433.
- Rafiee, A. and Hillestad, M. (2010). Optimal design and operation of a gas-to-liquid process. *Chemical Engineering Transactions*, 21, 1393–1398.
- Shah, M.D. (2011). Fault detection and diagnosis in nuclear power plant—a brief introduction. In *2011 Nirma University International Conference on Engineering*, 1–5. IEEE.
- Sorsa, T., Koivo, H.N., and Koivisto, H. (1991). Neural networks in process fault diagnosis. *IEEE Transactions on systems, man, and cybernetics*, 21(4), 815–825.
- Szargut, J. (2007). Egzergia. poradnik obliczania i stosowania. *Wydawnictwo Politechniki Slaskiej*.
- Venkatasubramanian, V., Rengaswamy, R., Yin, K., and Kavuri, S.N. (2003). A review of process fault detection and diagnosis: Part I: Quantitative model-based methods. *Computers & chemical engineering*, 27(3), 293–311.
- Watanabe, K. and Hirota, S. (1991). Incipient diagnosis of multiple faults in chemical processes via hierarchical artificial neural network. In *Proceedings IECON'91: 1991 International Conference on Industrial Electronics, Control and Instrumentation*, 1500–1505. IEEE.
- Wilson, D.R. and Martinez, T.R. (1997). Improved heterogeneous distance functions. *Journal of Artificial Intelligence Research*, 6, 1–34.
- Xie, L., Lin, X., and Zeng, J. (2013). Shrinking principal component analysis for enhanced process monitoring and fault isolation. *Industrial & Engineering Chemistry Research*, 52(49), 17475–17486.