Biomimetic and Advanced Control Structure Design with Real Time Optimization

Temitayo Bankole

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Recommended Citation

https://researchrepository.wvu.edu/etd/7159

This Dissertation is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Graduate Theses, Dissertations, and Problem Reports by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.
Biomimetic and Advanced Control Structure Design with Real Time Optimization

Temitayo Bankole

Dissertation submitted to the
Benjamin M. Statler College of Engineering and Mineral Resources
at West Virginia University

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in
Chemical Engineering

Debangsu Bhattacharyya, Ph.D., Chair
Richard Turton, PhD.
Fernando V. Lima, Ph.D.
Mario Perhinschi, Ph.D.
Urmila Diwekar, Ph.D.

Department of Chemical and Biomedical Engineering

Morgantown, West Virginia
2018

Keywords: Biomimetic control structure design, heuristic algorithms, system identification, optimization, advanced energy systems, real time optimization

Copyright 2018 Temitayo Bankole
Abstract

Biomimetic and Advanced Control Structure Design with Real Time Optimization

Temitayo Bankole

While numerous works exist in the area of control structure design from a holistic plantwide approach, this can be computationally intractable as process plants are typically characterized by a large number of variables which renders traditionally deployed process systems algorithms prohibitive. As parallelization and distributed computing become increasingly important and feasible, a method for structural analysis of plants which estimates connectivity strengths among various sub-processes making algorithms (including control structure design algorithms) amenable for distributed systems is proposed. In this thesis, analogy is drawn to the neuroscience literature where connectivity of neuronal population is established using data from magnetic resonance imaging. By using an input-state-output deterministic model for process systems and parameterizing this model to reflect connectivity and coupling, a Bayesian scheme is developed to estimate connectivity while incorporating priors. This connectivity is employed to subdivide an overall process into distinct islands for the purpose of control structure design. Consequently, for each island, a biomimetic multiagent approach stemming from the imitation of the central nervous system is deployed to coordinate and aggregate control structure design from each island for the overall process. This multiagent approach exploits coordination and communication found in nature to glean computational superiority. Additionally, this thesis addresses the controlled variable selection of a cyber-physical system for optimal economic operation. Finally, a real time optimization and scheduling of advanced energy power plants with CO₂ capture is developed and implemented.
Dedicated to

my Father, Abdul-Jelil, Bankole

Who showed me honor and how to be a man.

And to my daughter, Ariella,

To lay a foundation for great things to come

And to my darling wife Ayo,

Who has shown me love in its purest forms
Acknowledgments

I would like to express gratitude to National Energy Technology Laboratory (NETL) for their financial support during this period. The work was funded by the U.S DOE through the project titled “Development of Integrated Biomimetic Framework with Intelligent Monitoring, Cognition and Decision Capabilities for Control of Advanced Energy Plants” (Grant # DE-FE0012451).

Secondly, I would like to thank my advisor, Dr. Debangsu Bhattacharyya for his kind support, assistance and mentorship for the last couple of years as a graduate research assistant with the Advanced Process and Energy Systems Engineering Group. It has been an immensely challenging and rewarding journey. In addition, I would like to thank my committee members and the biomimetic research team, Dr. Richard Turton, Dr. Fernando Lima, Dr. Mario Perhinschi and Dr. Urmila Diwekar for their assistance and input throughout this journey. In addition, I would like to also mention researchers I also worked with during this period. Gaurav Mirlekar, Berhane Gebreslassie and Ghassan Al-Sinbol.

I would also like to thank colleagues and lab mates who have helped the stay warm and worthwhile. This includes Dustin, Jake, Chirag, Anca, Anderson, Josh, Pratik, Prokash, Yuan, Qiao and others. Finally, I would like to thank my wife for her enduring support.
# Table of Contents

Abstract .................................................................................................................................................. ii

Table of Contents .................................................................................................................................. v

List of Figures ......................................................................................................................................... viii

List of Tables ......................................................................................................................................... xi

Nomenclature ......................................................................................................................................... xii

Chapter 1 ............................................................................................................................................... 1

1 Introduction....................................................................................................................................... 1

1.1 Motivation and Background ........................................................................................................ 1

1.2 Biomimetic Approach to Control Structure Design .................................................................. 5

1.2.1 Self-Organizing, Biomimetic Control Structure Selection .................................................. 5

1.2.2 Multi-agent Optimization Framework .................................................................................. 7

1.3 Computation Tools ....................................................................................................................... 8

1.4 Thesis Organization ...................................................................................................................... 8

1.5 Research Output ........................................................................................................................... 9

1.5.1 Conference Presentation ....................................................................................................... 10

1.5.2 Conference Publication ......................................................................................................... 10

1.5.3 Journal Publications .............................................................................................................. 10

Chapter 2 ............................................................................................................................................... 11

2 Development of Algorithms for Biomimetic, Self-Organizing Control Structure Design... 11

2.1 Introduction................................................................................................................................... 11

2.2 Dynamic Causal Model.............................................................................................................. 15

2.2.1 Model Description ................................................................................................................. 17

2.3 Parameter Estimation with Bayesian Inferencing ..................................................................... 19

2.3.1 Model Integration .................................................................................................................. 19

2.3.2 Bayesian Inferencing ............................................................................................................ 21

2.3.3 Signal Estimation .................................................................................................................. 23

2.3.4 Expectation Maximization .................................................................................................... 24

2.4 Results & Discussions .................................................................................................................. 27

2.4.1 Toy Example ......................................................................................................................... 27

2.4.2 Van De Vusse Reactor with Separator .................................................................................. 29
Chapter 3
Development of Multi-agent Optimization Based Approach for Controlled Variable Selection

3.1 Background
3.2 Introduction
3.3 Approach
  3.3.1 A Priori Analysis
  3.3.2 Formulation of Loss Objective Function (Local Linear Exact Method)
  3.3.3 Controllability
  3.3.4 Posteriori Analysis
3.4 Decomposition of Process Architecture
3.5 Optimization Scheme
  3.5.1 Multiagent Optimization
3.6 Case Study: Acid gas removal (AGR) Unit
  3.6.1 Process Description
  3.6.2 Decomposition of AGR unit
  3.6.3 Identification of Candidate Sets of Controlled Variables
3.7 Conclusion

Chapter 4
Optimal Control Structure Design for Cyber-Physical Systems

4.1 Background
4.2 Introduction
4.3 Process Description
  4.3.1 Compressor/Turbine
  4.3.2 Heat Recuperation
  4.3.3 Bypass Valves
  4.3.4 Pressure Vessels
4.4 Controlled Variable Selection
  4.4.1 Gain Matrices
  4.4.2 Estimation of Implementation Error
List of Figures

Figure 1.1 Feedback policy with optimization and control layer with controlled variables $c_s$ as a combination of measured output variables $y_p, y_s$. Inputs, disturbances and noise denoted as $d, \omega$. ....3

Figure 1.2. Overview of biomimetic control approach to integrated energy plant ........................................... 6

Figure 2.1 Connectivity patterns represented by the DCM where each region denotes one or more unit operations. Connectivity between regions can be latent (black arrows) and/or induced (red dotted arrows) while external inputs are treated as stimulus-bound inputs (blue dash dotted)................. 17

Figure 2.2 Comparison of parameter estimation computation time................................................................. 28

Figure 2.3 Reactor separator set up for the Van De Vusse reactor ................................................................. 29

Figure 2.4 Concentration profile of P (left), Q (right) for a signal to noise ratio of 24 ............................. 31

Figure 2.5 Euclidean distance between noisy signal and signal estimate (left), Euclidean distance between true parameter vector and estimates (right), signal to noise ratio of 24. ................................. 32

Figure 2.6 Process flow configuration of the acid gas removal unit (modified from (Bhattacharyya et al., 2010)) ...........................................................

Figure 2.7 Normalized vapor phase concentration of $\text{CO}_2$ in stripper (left), liquid phase concentration of $\text{H}_2\text{S}$ (right) in the $\text{CO}_2$ absorber: true data (blue dash dot), estimates (solid red), noisy signal (black) ............................................................................................................................... 36

Figure 2.8 Normalized LP flash drum temperature (left) and $\text{H}_2\text{S}$ concentrator sump temperature (right).
......................................................................................................................................................... 36

Figure 2.9 Plot of fraction of strongly connected variables as threshold $\zeta$ increases......................... 39

Figure 3.2 The general framework of the multiagent platform................................................................. 54

Figure 3.3 An agent in MAOP framework (Gebreslassie and Diwekar, 2016). .......................... 55

Figure 3.4 Running time for BB (branch and bound), PBB (Parallelized branch and bound) and MAOP (Multiagent optimization) for the cases of 1, 2 and 3 islands respectively. .......................... 62

Figure 3.5 Economic loss - Average case (left) and Worst case (right) versus number of sections considered. ......................................................................................................................... 64

Figure 3.6 Controllability versus number of sections considered......................................................... 64

Figure 3.7 Average Loss (left) and controllability (right) as a function of variation in input syngas flow rate to the $\text{H}_2\text{S}$ absorber ......................................................................................................................... 68
Figure 5.9 Daily averaged electricity price (a) and demand (b) from January to March, 2015

Figure 5.10 Syngas flow rates (a) and CO$_2$ capture fraction set points (b) for the Scenario 1

Figure 5.11 (a) Syngas flow rates and (b) CO$_2$ capture fraction set points for Scenario 2

Figure 5.12 (a) Syngas flow rates and (b) CO$_2$ capture fraction set points for Scenario 3

Figure 5.13 Objective function (Profit) and cost of carbon capture for Scenario 1 (with RTO – black solid, without the RTO – blue dash dot)

Figure 5.14 Objective function (Profit) and cost of carbon capture for Scenario 2 (with RTO – black solid, without the RTO – blue dash dot)

Figure 5.15 Objective function (Profit) and cost of carbon capture for Scenario 3 (with RTO – black solid, without the RTO – blue dash dot)

Figure 5.16 Sensitivity of average carbon capture due to changes in the carbon tax γ($/ton CO_2$) (a)
Scenario 1; (b) Scenario 2; (c) Scenario 3
**List of Tables**

Table 2.1 Classification of strength for latent and induced connectivity .............................................. 27
Table 2.2 True and estimated parameter values for different algorithms ................................................. 28
Table 2.3 Parameters used for the Van de Vusse reactor separator ......................................................... 31
Table 2.4 Latent and Induced connectivity ................................................................................................. 33
Table 2.5 Classification of inputs and state variables for unit operations in CO\textsubscript{2} absorber  
flowsheet .................................................................................................................................................. 35
Table 2.6 Summary of induced connectivity table for the AGR unit* ....................................................... 38
Table 3.1 List of all candidate primary controlled variables ................................................................. 56
Table 3.2 Decomposition of the AGR unit into sections ........................................................................ 57
Table 3.3 List of manipulated variables and hardware for the AGR unit ............................................. 59
Table 3.4 Decomposition of the AGR unit into sections for CV selection ............................................. 59
Table 3.5 Parameters for the agents ......................................................................................................... 61
Table 3.6 Results of controlled variable selection using average loss economic and controllability  
objective function for different no of islands ......................................................................................... 63
Table 3.7 Indices of candidate controlled variables ................................................................................. 66
Table 4.1 List of all candidate controlled variables and their respective subsystem .......................... 76
Table 4.2 List of manipulated variables .................................................................................................. 77
Table 4.3 Candidate controlled variables in the hyper facility ........................................................... 77
Table 4.4 Parameters loaded into the SIMULINK workspace ............................................................... 80
Table 4.5 Estimates of noise variance for candidate controlled variable ........................................... 85
Table 4.6 Controlled variable Pareto set in descending order of optimality ........................................ 88
Table 5.1 Summary of objective function values and cost of carbon capture for all scenarios.. 118
<table>
<thead>
<tr>
<th>Nomenclature</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>Latent connectivity matrix</td>
</tr>
<tr>
<td>$\tilde{A}$</td>
<td>Concatenated latent connectivity matrix</td>
</tr>
<tr>
<td>$B$</td>
<td>Induced connectivity matrix</td>
</tr>
<tr>
<td>$\tilde{B}$</td>
<td>Concatenated induced connectivity matrix</td>
</tr>
<tr>
<td>$C$</td>
<td>Output matrix for underlying process</td>
</tr>
<tr>
<td>$\hat{C}$</td>
<td>Estimated output matrix for underlying process</td>
</tr>
<tr>
<td>$\bar{C}$</td>
<td>Output matrix for prediction model</td>
</tr>
<tr>
<td>$C_e$</td>
<td>Estimated error covariance of process model</td>
</tr>
<tr>
<td>$C_Y$</td>
<td>Covariance of measurements</td>
</tr>
<tr>
<td>$d$</td>
<td>Disturbances</td>
</tr>
<tr>
<td>$D_e$</td>
<td>Scaling matrix for output variables</td>
</tr>
<tr>
<td>$D_u$</td>
<td>Scaling matrix for input variables</td>
</tr>
<tr>
<td>$\mathbb{E}(\cdot)$</td>
<td>Expectation</td>
</tr>
<tr>
<td>$e$</td>
<td>Stochastic error term employed in prediction</td>
</tr>
<tr>
<td>$E$</td>
<td>Electricity prices</td>
</tr>
<tr>
<td>$f$</td>
<td>Revenue generation term in the objective function</td>
</tr>
<tr>
<td>$F(\cdot)$</td>
<td>Nonlinear state space model</td>
</tr>
<tr>
<td>$F_{CO_2}$</td>
<td>$CO_2$ flow rate</td>
</tr>
<tr>
<td>$F_{out}$</td>
<td>Flow rate of syngas/fluegas post carbon capture</td>
</tr>
<tr>
<td>$F_s$</td>
<td>Syngas flow rate</td>
</tr>
<tr>
<td>$G$</td>
<td>Steady state gain matrix</td>
</tr>
<tr>
<td>$G_d$</td>
<td>Steady state disturbance gain matrix</td>
</tr>
<tr>
<td>$G^\gamma$</td>
<td>Full steady state gain matrix</td>
</tr>
<tr>
<td>$G^\gamma_d$</td>
<td>Full state steady state disturbance gain matrix</td>
</tr>
<tr>
<td>$\hat{G}$</td>
<td>Scaled gain matrix</td>
</tr>
<tr>
<td>$h(\cdot)$</td>
<td>Integrated process model</td>
</tr>
<tr>
<td>$H$</td>
<td>Second order nonlinear latent connectivity</td>
</tr>
<tr>
<td>$i$</td>
<td>Current time instant</td>
</tr>
<tr>
<td>$J$</td>
<td>Objective function</td>
</tr>
<tr>
<td>$J(\cdot)$</td>
<td>Penalty cost function on $CO_2$ capture</td>
</tr>
<tr>
<td>$J_c$</td>
<td>Controllability objective function</td>
</tr>
<tr>
<td>$J_{control}$</td>
<td>Objective function for control complexity optimization</td>
</tr>
<tr>
<td>$k$</td>
<td>Time step index within each horizon</td>
</tr>
</tbody>
</table>
\( l \) Stage cost function

\( L \) Loss

\( LHV \) Lower heating value

\( m \) Number of steps in near horizon

\( N \) Number of steps in the base time at an optimization run

\( N_d \) Dimensionality of disturbances

\( N_u \) Dimensionality of inputs

\( N_y \) Dimensionality of outputs

\( p(u_k) \) Cost function of producing syngas and CO2 capture

\( P \) Logical set of vectors

\( q \) State variables of the prediction model

\( Q \) Observability Gramian

\( \dot{Q} \) Heat rate

\( T \) Frequency of scheduler run (one hour)

\( T_s \) Sampling frequency of lower level controller

\( u \) Degrees of freedom of the scheduler, i.e. set points passed to the supervisory control layer

\( \bar{u} \) Control input sequence from time instant \( i \) to end of base period \( N \)

\( u_{1,min} \) Minimum carbon capture

\( u_{2,max} \) Maximum carbon capture

\( \Delta u_{1,min}, \Delta u_{2,max} \) Max uprate and downrate CO2

\( u_{2,min} \) Minimum syngas flow rate

\( u_{2,max} \) Maximum syngas flow rate

\( \Delta u_{2,min}, \Delta u_{2,max} \) Max uprate and downrate syngas

\( v \) Control inputs at the supervisory control layer

\( V \) Optimization cost function for the scheduler

\( w_1 \) Weighting factor for near horizon

\( w_2 \) Weighting factor for far horizon

\( W \) Power

\( x \) State variables for underlying power plant process.

\( y \) CVs at the supervisory control layer

\( z_c \) CO2 capture fraction
**Greek letters**

- \( \alpha \) CO\(_2\) capture target
- \( \theta \) Parameters
- \( \omega \) noise
- \( \sigma \) Singular values
- \( \delta \) Summation of past errors in carbon capture
- \( \eta \) Efficiency
- \( \lambda \) hyperparameters
- \( \zeta \) Connectivity threshold
- \( \xi \) Upper limits on CV selection
- \( \chi \) Upper limit on time delay
- \( \gamma \) Carbon tax
- \( \epsilon \) Slack variable for carbon tax penalty term
- \( \Psi \) Weights on output in MPC formulation
- \( \Phi \) Weights on input movement suppression

**Abbreviation**

- AGR Acid gas removal unit
- ARIMA Autoregressive integrated moving average
- BA Bleed air
- CA Cold air
- CV Controlled variables
- DCT Discrete cosine transform
- DCM Dynamic causal model
- DV Disturbance variables
- GCV Generalized cross validation
- GT Gas turbine
- HA Hot air
- HP High pressure
- HX Heat exchanger
- HyPer Hybrid performance
- IGCC Integrated gasification combined cycle
- ISE Integral square error
- LHV Lower heating value
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMPC</td>
<td>Linear model predictive control</td>
</tr>
<tr>
<td>LP</td>
<td>Low pressure</td>
</tr>
<tr>
<td>MAOP</td>
<td>Multiagent optimization</td>
</tr>
<tr>
<td>MPC</td>
<td>Model predictive control</td>
</tr>
<tr>
<td>MV</td>
<td>Manipulated variables</td>
</tr>
<tr>
<td>nom</td>
<td>Nominal</td>
</tr>
<tr>
<td>PI</td>
<td>Proportional integral</td>
</tr>
<tr>
<td>RSS</td>
<td>Residual sum of squares</td>
</tr>
<tr>
<td>SOFC</td>
<td>Solid oxide fuel cell</td>
</tr>
</tbody>
</table>
Chapter 1

1 Introduction

1.1 Motivation and Background

According to the highlights of the world energy outlook, energy demand is projected to increase by 37% by 2040 (EIA, 2013). As demand continues to grow, the need for clean energy is increasingly becoming important. Constrained by ever tightening environmental regulations and demand for increased plant availability, high efficiency and profitability remains a crucial requirement for power plants. Therefore, operations of energy plants need to be profitable, agile and flexible while maintaining maximum efficiency. This necessitates advanced optimal strategy for operations. A crucial part of process operations is the control structure design of process plant i.e. the selection of the optimal controlled variables (CVs). In this research, mathematical tools are leveraged for solving aforementioned challenging problems associated with the optimal CV selection using biologically inspired techniques. In addition, this work also focuses on optimization and scheduling of set points of pertinent controlled variables for an energy plant. This is necessary as changes in disturbances necessitate changes in set points of CVs therefore periodic optimization must be performed in the face of stochastic predictions of disturbances to calculate and pass these set points to the controller (supervisory control layer). Together, these constitute a necessary and important part of the optimal requirements of energy plants in the near future.

A chemical process plant is operated with an objective that is desired to be optimized. To achieve the optimal operation, a number of variables needs to be measured, manipulated and controlled. Traditionally, previous works in open literature have based the selection of controlled variables on heuristics lacking a methodical approach (Fisher et al., 1985). The earliest works include that of (Murthy Konda et al., 2005) where an integrated framework of heuristics and simulation are provided as a means for plantwide control. This was an improvement on the works of (Luyben et al., 1997) where a nine step heuristic based method was outlined for complex processes consisting of various process units. (Morari et al., 1980) developed mathematical measures within the framework of multilevel optimization theory for decomposition and partitioning of processes for the purpose of control. These included studies of the effect of controlled variable selection on plant
operation. Other works in literature have also addressed controlled variable selection in one form or the other (Narraway et al., 1991, Arbel et al., 1996, Rijnsdorp, 1991, Zheng et al., 1999). Recently, a systematic approach to optimal CV selection by considering an economic loss function has been proposed (Skogestad, 2004). However, the CVs selected by considering only economic criterion may need to poor controllability. To circumvent this issues, some measures of controllability and control performance were included in the CV selection process by Jones et al. (2014). Jones et al. (2014) proposed a three-stage procedure (a priori, optimization, posteriori) for selection of primary CVs for processes that operates with a hierarchical control structure like shown in Fig. 1.1. At the upper layer, a real time optimizer (RTO) that typically use a steady-state model, periodically (typically minutes or hours) updates the setpoint for the primary CVs, which in turn, updates the setpoints of the secondary CVs.

During the a priori analysis, manipulated variables (MVs) and disturbance variables (DVs) are identified and a list of the candidate CVs for the primary control layer is generated, but a significant number of candidate variables that do not have acceptable gain for servo control and disturbance rejection is eliminated. The remaining CVs are further analyzed at the second stage. This is the most important step, where first, the process and operational constraints that are active under the desired design and off-design conditions are identified. These active constraints are selected as primary CVs. A number of additional CVs are also selected depending on the additional degrees of freedom. In the final stage, a posteriori analysis is performed for the CV sets selected at the 2nd stage. This step is necessary since a linear process model is used in the 2nd stage. Therefore, the economic and control performances of the CV sets from the second stage are evaluated under off-design conditions by using a nonlinear process model.

---

1 By controllability, we mean ease of control and not necessarily being controllable as defined in classical control theory for linear time invariant systems.
Figure 1.1 Feedback policy with optimization and control layer with controlled variables $c_s$ as a combination of measured output variables $y_p,y_s$. Inputs, disturbances and noise denoted as $d,\omega$.

The most time-consuming step is the second step due to the combinatorial nature of the optimization problem. Even after prescreening of the candidate CVs, there can be large number of CV sets that needs to be evaluated during the 2nd stage. This combinatorial optimization problem can easily explode with the increase in the number of potential CVs that can be often correlated with the plant size/complexity. For a small plant where one has to choose 10 CVs from 80 for example, the combinatorial demand is $\binom{80}{10} = 1.6 \times 10^{12}$. Typically, branch and bound (BB) optimization methods have been used for solving the optimization problem in the 2nd stage (Cao and Kariwala, 2008, Kariwala and Cao, 2009, Jones et al., 2014). More recently convex optimization (Yelchuru et al., 2010) has been proposed for controlled variable selection with constraints. However, solving the optimization problem where trillions of combinations need to be evaluated can be computationally prohibitive and therefore is not suitable if re-selection of the CV sets needs to be done often. It can be noted that re-selection of the CV sets is desired when any of the following things change with respect to the nominal operation - operational objective, update...
in the list of CVs, MVs or DVs or their bounds, or the underlying process models. If one or more new equipment item(s) is/are added or removed or the configuration of the process units are changed, then not only the list of CVs, MVs or DVs or their bounds needs to be changed, but the underlying process model also needs to be updated. Example of one such process is the cyber-physical processes where the cyber-component of the process can be readily modified, replaced or the process configuration can be readily changed. Operational objective of the newly configured process is likely to change as well. It can also be noted that in chemical plants, change in the operational objective is also common. A plant can operate to maximize profit or maximize production or maximize yield or minimize utility consumption, for example. Thus it is desired that the CV selection process be executed faster than the current state-of-the-art. For this reason, a number of features in biological systems would be adapted for use.

Self-organization, distributed intelligence, adaptability, intelligent monitoring, and decision capabilities are some of the characteristics of the biological world that can be effectively utilized in the optimal control structure design of plants. An example of a distinguishing feature of biological systems is information sharing and cooperation. The proposition in this research is that the computational demand of CV selection can be reduced drastically if a process is considered as different sections rather than holistically. Thus the CV selection can be performed independently on each section and the results can be aggregated thus mimicking cooperation, divide and conquer found in biological systems. Additionally, another strategy presented in this work is to employ metaheuristic biologically inspired optimization techniques as opposed to branch and bound. It is proven that these two strategies would improve the computational time thus energy plants can afford to deploy CV selection more often than is currently realizable in the industry.

In addition to selecting CVs, optimality of plant operations depend on CV set points as they are results of an optimization performed at a nominal point, thus these set points become sub-optimal once disturbances change from nominal operating points. Therefore this optimization must be periodically performed to obtain new set points as the process navigates from one operating point to another, this necessitates the mathematical framework of real time optimization in chapter 5.

Consequently, this thesis presents the theoretical development and practical implementation of biologically inspired techniques for optimal control structure design of advanced energy plants. The aim of this is to improve flexibility, optimality and efficiency of advanced energy plants now
and in the near future. In addition to this, real time optimization in the presence of stochastic disturbances for the purpose of maximizing economic profit while maintaining environmental emission standards is presented. This portion however employs conventional rather than biomimetic approach for its development.

1.2 Biomimetic Approach to Control Structure Design

This research is part of an overall biomimetic approach to control structure and controller design for an advanced energy plant as shown in Fig. 1.2. A self-organizing, biomimetic control structure selection process dynamically adapts the controlled variables for maximizing the plant profit without violating constraints. The controller design process then accepts these sets of controlled variables and designs centralized/decentralized controllers that exploit the rule of pursuit present in ant colonies. To reject the modeled and unmodeled disturbances, an intelligent system monitors the process and adapts the control actions by infusing cognition and decision capabilities.

1.2.1 Self-Organizing, Biomimetic Control Structure Selection

The overall scheme in Fig 1.2 is as follows: objectives and disturbances are passed into the system, in order to meet this objective, self-optimizing CVs must be selected. The term self-optimizing implies that operating the process plant while keeping the CVs constant at predetermined set points will result in an acceptable loss (Skogestad, 2004). Loss is defined as the difference in the objective function values between the optimal cases as compared with when CVs are kept at the constant setpoint provided by the RTO at the upper layer. This is infeasible to solve in real time or in every couple of minutes if all the candidate variables are considered. Here, to reduce the size of the problem and still achieve self-optimizing control, the function of the cortical areas of human brain is imitated. Thus process data is then used to establish the partitioning/decomposition of the system into various sections/islands. This is seen in Fig. 1.2 as regions 1 through 5 (R1-R5).

To perform this decomposition, it is required to determine the specific pattern and intensity of connections in response to the stimulation. Therefore units with strong couplings and connectivity would be considered together during portioning and otherwise. To this end, first a process model referred to as the Dynamic Causal Model (DCM) is utilized to establish this connectivity strength. In neuroscience, the brain is considered to be a deterministic input-state-output process and an analogous connectivity estimation approach is used to understand the self-organization of the
cortical areas of the brain. The inputs in DCM are conventional stimulus functions that are analogous to manipulated variables in process control.

Upon decomposition, each section/island would have its CVs established. These CVs (measurable and observable) are then aggregated together and passed to the supervisory control layer (beyond the scope of this thesis) where controller design is performed. During process operation, intelligent monitoring of the process is performed to establish when the process departs into abnormal conditions (Al-Sinbol, 2013). As the process operation moves from one operating point to another, or when objective function changes, it may be necessary to repeat the process for the purpose of reorganizing the decomposition and consequently CV selection to meet the new demands of the process. This selection procedure requires solution of a mixed integer nonlinear programming (MINLP) problem through the multi-agent optimization framework that mimics the CNS.

It should be noted however, that change in CV for a plant during operation is rather futuristic and philosophical for now as most plants can only afford this during start up after a period of shut down. When energy plants are completely automatic, this may very well be applicable. Therefore the utility of the propositions and methodologies developed in this work would find application in a near or completely autonomous plant.

Figure 1.2. Overview of biomimetic control approach to integrated energy plant
1.2.2 Multi-agent Optimization Framework

The control structure design problems for biomimetic control of power plants are expected to be nonconvex. The possibility of obtaining local minima or maxima is very high in such problems. Some of the biomimetic optimization techniques such as genetic algorithms, ant colony optimization (and simulated annealing) show higher probability of obtaining global solutions. However, these techniques can be computationally intensive. For each island, one MINLP problem (for control structure selection) and several NLP problems (one for each controller present in an island) need to be solved. This can result in computational intractability for large scale systems. For the bilaterian animals, the CNS coordinates the activities of the entire body in real-time in an optimal manner. To achieve the similar functionalities as the CNS, a multi-agent optimization framework will be developed in this thesis. The multi-agent optimization framework provides a way of combining various algorithms in one platform and exploits the strengths that each one of them possesses. Such an approach avoids the problem of getting stuck in local optima as well as reduces the computational burden. In process systems engineering, agent-based systems are proposed for conceptual design (Chonghun et al., 1995), supply chain management (Julka et al., 2002, Mele et al., 2007), and controller design (Tetiker et al., 2008, Tatara et al., 2005).

However, for multi-agent optimization of process systems engineering problems, very few articles have appeared and most of them are restricted to small scale problems (Sirola et al., 2003). In this work, the multi-agent optimization framework was implemented for a large-scale, real-world problem. The flow of data and control structure will be similar to Sirola et al (Sirola et al., 2003). However, the agents will be designed differently for solving large scale optimization problems. There will be three autonomous agents consisting of transient programs that run independently on various machines. The three agents include the efficient simulated annealing agent (ESA), efficient genetic algorithm agent (EGA), efficient ant colony agent (EAC). Since it is expected that the number of agents will be more than the machines available to run them on, the central executive routine will schedule the agent runs based on the waiting time and probability function assigned to each agent depending on previous successes. For example, GA agents can find global optimum with lesser computational effort than the other agents for certain classes of problems; so the central agent might give higher probability of success to the EGA agent for these types of problems. This scheduling algorithm is based on efficient sampling techniques developed by Diwekar’s group over the last decade (Diwekar and Ulas, 2007).
The agent provides search regions to the common memory space, to be explored by the optimizers based on previous clustering agent’s work. The optimizing agents use different algorithms to solve the same problem. These agents will be running on different machines. However, every optimization agent will perform small number of iterations than typically required for reaching optimal solution. This increases the frequency at which those agents can communicate and collaborate with the other agents in the system by requiring them to post their solutions and then reinitialize more frequently. The ESA, EGA, and EAC agents are based on new efficient heuristic based algorithms called Efficient Simulated Annealing, Efficient Genetic Algorithm, and Efficient Ant Colony algorithm. The first two algorithms (ESA and EGA) have been developed by considering the k-dimensional uniformity of a quasi-random number generator based on Hammersley Sequence Sampling (Kalagnanam and Diwekar, 1997) developed in Dr. Diwekar’s group. The multi-agent optimization framework proposed above to handle both MINLP and NLP problems is a unique framework that is developed for the first time for control of power systems.

1.3 Computation Tools

MATLAB®, a computing environment developed by MathWorks®, is one of the main engineering software used in this project for modeling and optimization. Another major software used for the completion of this project is DYNSIM. The Dynsim - Matlab engine Link is an interface for including Matlab-based models and controllers in a DYNSIM dynamic simulation model using the OPC data access protocol. The engine link is capable of handling several scenarios including, solution unavailability, solution impossibility, and data communication loss. The link was developed by Schneider Electric for West Virginia University.

1.4 Thesis Organization

Chapter 1 presents the background and motivation. Chapter 2 focuses on the first biologically inspired algorithm which is a decomposition algorithm stemming from the analogy of the human cortical brain. In this chapter, a process plant is viewed as a coordinated system of different sections/islands with connectivity existing amongst them. This is done through a dynamic causal model (DCM). This connectivity is thought to be modelled after neuronal connections found in the human brain. Borrowing from the self-organization of the human brain in neuroscience, this task aims at developing the framework for distributed intelligence and computing for energy
plants. Connectivity information between different processes of the energy plant are garnered through probabilistic network methods. This result can be used to aggregate strongly connected islands together for the purpose of deploying algorithms such as the control structure design.

Chapter 3 exploits the results of Chapter 2 for the purpose of control structure design. In particular, the use of a coordinated multiagent platform is discussed which employs exploitation and exploration to achieve faster convergence on optimization problems as opposed to conventional methods. The strength of coupling between various controlled variables will be evaluated by using the DCM so that different islands with strongly coupled controlled variables can be identified. Each island will form an independent sub-problem. This will be followed by a development of a multi-agent optimization (MAOP) for each island to select the controlled variable using the results from the DCM. This multi-agent system will solve a mixed integer nonlinear programming (MINLP) problem by mimicking the distributed intelligence of the central nervous system (CNS). The information obtained therein about input-state-output interrelations available from the DCM and the MAOP can then be passed on to control configuration design.

In Chapter 4, the methods of chapter three are extended to a cyber physical system with virtual components, the biomimetic CV selection is deployed to a fuel cell gas turbine hybrid system. This system poses unique characteristics which render multiagent coordination attractive to employ.

In Chapter 5, a real time optimization algorithm is proposed based on economic optimality. Production and carbon capture are scheduled based on stochastic predictions of future electricity demands and electricity prices while meeting environmental regulations.

In Chapter 6, recommendations and future research directions are provided.

1.5 Research Output

The contributions of this research includes:

1. Algorithmic development of connectivity estimation with a second order nonlinear model
2. Decomposition and partitioning algorithm of process plants based on structural connectivity
3. Partition based CV selection predicated on inferred structural connectivity
4. Use of multiagent metaheuristic algorithms for CV selection
5. Real time optimization of an energy plant with CO₂ capture
6. Lyapunov stability of the mathematical formulation of RTO of energy plant with CO₂ capture
The contributions of this research in publications and presentations are as follows:

1.5.1 Conference Presentation


1.5.2 Conference Publication


1.5.3 Journal Publications


Chapter 2

2 Development of Algorithms for Biomimetic, Self-Organizing Control Structure Design

2.1 Introduction

Many process systems engineering tools at the heart of optimization and control require the solution of large scale problems which demand significant computational expense (El-Beltagy et al., 1999). Recent advances in development of theoretical tools in control and optimization together with the state of the art computational power and available software have further opened up immense possibilities. In spite of the increased performance and efficiency of computing speed and power, it is still infeasible to solve large-scale process optimization problems especially when the application is intended for online deployment or fast computation of the solution is desired. Examples of such large-scale optimization problems include, but are not limited to: various dynamic optimization problems that are solved for obtaining optimal control trajectory, reconciling dynamic data, or for obtaining optimal estimates of time-varying parameters, online adaptation of process models etc. If the underlying problem is combinatorial in nature, then the optimization problem can be computationally prohibitive even for moderate-sized plants. An example of such a large-scale combinatorial problem is the controlled variable (CV) selection problem (Jones et al, 2014). It should be noted that while CV selection is done heuristically or by off-line evaluations that are rarely revisited, the CV sets can be sub-optimal if the control objective of the process changes or the process model or operating constraints change considerably. One example of such processes is a cyber-physical system where the control objective can considerably change over a period of time or components can be readily added/modified/removed changing the underlying process model and operating constraints. Thus fast (not necessarily online) selection of updated optimal CV sets will be highly desired. It can be noted that the search for optimal CV sets involves systematic evaluation of an objective function for large number of candidate sets. For highly nonlinear plants, solution of the underlying optimization problem for each candidate set is difficult and computationally demanding. One approach to solving such large scale optimization problems is to use a ‘divide and conquer’ approach where a large process can be decomposed into smaller sub-processes. Such decomposition can be accomplished using inferred structural
connectivity information among various sub-processes. Then, only strongly connected sub-processes can be considered together for computation. Therefore, a methodology that can identify and detect dynamic changes in the connectivity among various sub-processes is desired.

Analysis of connectivity of chemical processes have long been researched in the open literature mainly from the perspective of fault detection and diagnosis. In those works, typically, a system is represented by using directed input and output arcs or signed digraphs (SDGs). These directed arcs represent causality. From these diagraphs, subsets of strongly connected components and maximal strongly connected component can be deduced. Strongly connected components in these sense are a combination of nodes that can be reached from every other node within the subset while a maximal strongly connected subset is a strongly connected subset with no input arcs (Iri et al., 1979). In (Emmerich et al., 2001) process plants have been modelled as structured graphs, a type of directed graphs. Unit operations are represented by vertices while edges represent streams from the outlet of a unit operation to the inlet of another. Causality is modelled in these graphs by using inlet and outlet connectors. An excellent review of various works in this area can be found in (Venkatasubramanian et al., 2003). The DGs and SDGs have been widely utilized from the perspective of fault propagation where steady state or incipient changes in the process variables are utilized for obtaining the connectivity information rather than considering a state-space model. Therefore, dynamic change in the connectivity due to dynamics of inputs cannot be inferred from the DGs and SDGs. Another drawback of the DG or SDG-based approaches are that a binary information (a value of ‘1’ if two nodes should be connected, ‘0’ otherwise) is obtained about the connectivity, but a quantitative measure of the relative strength in connectivity between various nodes is not available. A quantitative measure of the connectivity strength can be helpful in determining how to decompose processes.

Connectivity estimation is also important for the purpose of control structure selection. Input-output interaction can be quantified using participation matrices (PM)(Conley and Salgado, 2000), Hankel Interaction Index Array (HIIA) (Wittenmark and Salgado, 2002) and the $\Sigma_2$ measure (Birk and Medvedev, 2003). An estimation of interaction parameters for high order Vector ARX (VARX) has also been proposed (Carvalho Bittencourt, 2016) . These interaction parameters allow control pairing with superior performance compared to the relative gain array (RGA). A number of these methods have been compared in terms of computational complexity for control structure selection
Similarly, methods for interaction analysis using weighted graphs for control structure selection have been reported (Arranz and Birk, 2012).

In light of identifying connectivity in the field of neuroscience, two prominent methods used for estimating connectivity include Granger causality (or G-causality) and transfer entropy approaches. These approaches enjoy wide use in literature and typically employ autoregressive models. It has been reported that the Granger causality based approaches might perform poorly in comparison to other methods including partial correlations, mutual information, coherence, generalized synchrony and Bayesian networks as the measurement noise can reverse the estimation of causality direction (Smith et al., 2011). Extension of the Granger causality to nonlinear process systems where the nonlinearity can stem forth due to interaction between the input and state variables as well as due to interaction among the state variables is not straightforward. The use of transfer entropy for measuring process connectivity for fault diagnosis including process connectivity has been reported (Landman and Jämsä-Jounela, 2016). These two methods are typically applied when the variables are assumed to be Gaussian (Barnett et al., 2009). Other approaches include model-driven approaches generally known as structural equation modeling where specific model structures can be employed (Kline, 2015). The model-based approaches have been widely used in the area of economics, social sciences, and neuroscience, to name a few. For example, in the area of neuroscience, a modified direct transfer function model has been proposed where a multivariate auto-regressive model is converted to frequency domain and a partial coherence metric multiplied by the direct transfer function is used in quantifying connectivity (Korzeniewska et al., 2003). Excellent reviews of various methods for determining structural connectivity can be found in (Friston, 2011) for neuroscience and (Yang et al., 2014) for process plants.

Since the objective of the current work is to decompose the process model based on structural connectivity information, the connectivity measures are constrained by the physical connectivity of the process equipment items and the mass and heat exchange between them. Thus, a structural equation modeling approach is required. Due to the very nature of the typical chemical process systems and specific to the desired outcome of the current work, the candidate model should be nonlinear and should capture the nonlinearity due to interactions between states and inputs as well as interactions among state variables. Furthermore, it is desired that the stochastic parameters that quantify the structural connectivity be estimated for a non-Gaussian system. To the best of our
knowledge, there is no work in the current literature on quantifying structural connectivity of non-Gaussian chemical process systems characterized by bilinear models incorporating interactions between state and input variables as well as interactions between state variables.

For obtaining quantitative measure of connectivity strength and its dynamics, a dynamic model representing the process is desired where the model parameters would represent connectivity. Typical approach to candidate model selection for an intended application starts off with the qualitative measure of the system description, where the key features describing the system is identified. This step is usually referred to as the structural identification (Kay et al., 2000, Bradley and Stolle, 1996). In this thesis, a second order nonlinear model is considered as a candidate model. Its integration and use for connectivity estimation are original works of the author. The particular form of second order model considered in this work (Bankole and Bhattacharyya, 2018) is an extension of bilinear models found in the literature. In this model, the bilinear terms represent interactions between states and inputs as well as interactions among state variables. This is crucial in chemical engineering systems where exogenous inputs such as feed flowrates, temperature, and compositions have strong effects on states such as concentrations and temperatures. Bilinear models have been used in the field of neuroscience for the modelling of interactions amongst neuronal populations at a cortical and subcortical level (Friston et al., 2003). Using magnetic resonance imaging, evoked brain responses are used to characterize plausible models by making inferences about the coupling of several brain regions and the modulation of these couplings by experimentally designed inputs. By treating the brain as a deterministic input-state-output system, effective connectivity is parameterized as a function of couplings amongst unobserved neuronal states. However, the inferences are contingent upon assumptions about model structure. This is inevitable as concrete information about the architecture of the neuronal connectivity is unknown and can at best be surmised. Nevertheless, the utility of the dynamic causal model is grounded on its use as an exploratory means for model selection amongst several models (Will et al., 2004).

Once the candidate model is determined, model parameters must be estimated using one of several methods such as the minimum square error method (Ljung, 1987), maximum a posteriori method (Nelles, 2013), etc. A number of authors has used gradient-based methods for maximization of the likelihood function conditioned on the defined statistics of the observed data (Fnaiech and Ljung, 1987, GAB and Subba Rao, 1984, Verdult, 2002, Verdult et al., 2002). Schön et al. (Schön et al., 2011) provides a rigorous approach for system identification for a general class of discrete time
nonlinear systems with unknown parameters. Here, an expectation maximization algorithm is also used in this work. This algorithm seeks to maximize the likelihood of parameters conditioned on observed data and predefined priors while also maximizing the unknown statistics of the model error.

The proposed method developed in this work can perform satisfactorily in presence of noisy data of known variance. The approach is applied to two case studies. One case study pertains to a small sized system for which the exact model is available and therefore the connectivity information is known and therefore serves as a validation of the approach. Another case study pertains to a process of considerable size for which the exact model is not known but connectivity information can be inferred from process heuristics and therefore serves to validate the approach for high-dimensional system. The utility of the algorithm developed in this work is to decompose a process into smaller sub-processes such that the optimization/computational problems that are desired to be solved online or reasonably fast can be solved independently and/or in parallel for these sub-processes.

2.2 Dynamic Causal Model

To obtain the connectivity information in a process, it is sought to model output response with a parametric model whose parameters provide insight into the different classifications of the connectedness of the variables of interest. Candidate models for obtaining connectivity information should be such that they: (1) can be used for obtaining structural connectivity information, both inherent as well as those induced by internal changes and external disturbances, (2) can be developed using available simulation data, and (3) are reasonably simple so that they can be solved in real-time applications. Bilinear Dynamic Causal Models (DCM) are potential candidates that can satisfy these requirements. These models have been reported to have the capability to capture causal effects of stimulus-free contextual inputs as well as stimulus-bound perturbations on the connectivity among the cortical/sub-cortical areas in the brain (Friston et al., 2003).

In neuronal networks, the DCM is employed to observe unilateral and bilateral connectivity between different regions of the brain to infer structural changes modified by experimentally designed inputs. These models are estimated using a probabilistic approach where inferences about connectivity are evidenced by the posteriors. In process systems, the physical connectivity of the
process is usually known, thus as opposed to obtaining different model evidences to infer structural coupling, the objective here is to obtain quantitative information about the strength of connections between different process unit operations and exploit these connection for the decomposition purpose. Dynamic changes in the strengths of connection are evidenced by the estimated parameters of the employed model.

In using the DCM for the analysis of the functional integration of system’s dynamics, various regions (here representing unit operations) and the variables of interest must be identified. These comprise a set of \( i = 1, \ldots, N \) unit operations within the system. A typical process consisting of different process regions is shown in Fig. 2.1. Each region or node is denoted by \( R_i \). These regions would typically represent one or more equipment items or a subset of variables. Two types of connectivity exist between these regions- latent connectivity and induced connectivity. Latent connectivity denotes the intrinsic connectivity that exists among these regions. In Fig. 2.1, these are represented by black solid arrows representing the direction of influence or causality. These could be bidirectional or unidirectional, thus while the black arrow from \( R_1 \) to \( R_2 \) represents a bilateral forward and backward latent connectivity, the latent connectivity from \( R_4 \) to \( R_3 \) denotes unidirectional backward connectivity. In this notation, the numbering system of the regions is assumed to increase as one traverses the process downstream. Therefore a region \( R_i \) could represent a reactor while \( R_{i+1} \) represents a separator downstream the reactor. Thus backward connectivity could typically correspond to a recycle loop between two nodes. It is noteworthy that self-connectivity is omitted in the diagrammatic representation but is encoded in the latent connectivity matrix as discussed in Subsection 2.2.1. Induced connectivity (red dotted arrows) denotes connectivity that are modulated by the input, this type of connectivity can be seen from region \( R_2 \) to region \( R_3 \) where both a forward and a backward induced connectivity are activated by inputs. Contrarily, only a backward induced connectivity exists from \( R_2 \) to \( R_1 \). In addition, there exists extrinsic effect of inputs (blue dash dotted arrow) as can be observed in region \( R_1, R_2 \) and \( R_3 \). These effects denote the local effects of inputs on the process nodes.
2.2.1 Model Description

Consider a general nonlinear system described by a set of differential and algebraic equations as given below:

\[
\dot{x} = F(x, u, \theta) + v 
\]

\[
y = g(x) + \omega 
\]

Here, \(x, u, y, \theta\) represents state space vector, inputs, measurements and parameters respectively. The dimensionality of the variables is given by \(x \in \mathbb{R}^{N_x}, u \in \mathbb{R}^{N_u}, y \in \mathbb{R}^{N_y}\). Let \((x, y, u) = (0, 0, 0)\) denote steady state. It is assumed that the output signal is corrupted with additive measurement noise \(\omega\) while \(v\) denotes noise in the model (i.e. unmodeled, unknown, and/or inaccurate physics).

Approximating the general nonlinear model in Eq. (2.1) with a second order nonlinear model (truncated Taylor series expansion) given as:

\[
\dot{x} \approx Ax + \sum u_j B_j^j x + Cu + \text{diag}(x) H x 
\]

Where:

\[
A = \left. \frac{\partial F}{\partial x} \right|_{ss} 
\]
\[ B^j = \frac{\partial^2 F}{\partial x \partial u_j} \bigg|_{ss} \]  
\[ C = \frac{\partial F}{\partial u} \bigg|_{ss} \]  
\[ H = \frac{\partial^2 F}{\partial x^2} \bigg|_{ss} \]  

Here, diag(\(x\)) is a diagonal matrix where leading diagonal elements form the state space vector \(x\) such that the sequence of entries on an arbitrary row \(i\) of this matrix can be written as \(\{x_j \delta_{ij}\}_{j=0}^{N_x}\), where \(\delta\) is the Kronecker delta, i.e. \(\delta_{ij} = 1\) if \(i = j\), \(\delta_{ij} = 0\) if \(i \neq j\). The matrices \(\{A, B, C, H\}\) can be obtained from the nonlinear model given in Eq. (2.1) with appropriate differentials of \(F(x, u, \theta)\) evaluated at steady state (‘ss’) with respect to state, state and input, and input variables, respectively. Beyond mathematical significance, these matrices also give insight into the connectivity of the system. The Jacobian matrix \(A\) represents first order coupling among state variables. Elements of \(A\) represent hidden couplings of state variables devoid of exogenous inputs. In other words, these elements represent a fundamental structure of the system under consideration. In neuroscience, an analogy exists between the Jacobian matrix and the latent connectivity between neurons, which is an intrinsic coupling unmodulated by experimentally designed inputs. The eigenvalues of the matrix \(A\) designate the neuronal time constants of the brain regions and are assumed to be the same for all regions. For process systems, however, they represent the argument of the matrix exponential for the zero input case of a continuous time linear state space model. Similarly the matrices \(B^j\) signify couplings due to the effect of the \(j^{th}\) input \((u_j)\). These embody the interaction between state space variables subject to the influence of inputs. A distinct \(B^j\) matrix exists for each input and these are referred to as the induced connectivity matrices. The matrix \(C\) characterizes the effect of external inputs on the state variables. Lastly, the matrix \(H\) represents the Hessian matrix with respect to state variables. The last term in Eq. (2.3) encodes nonlinear latent connectivity between state variables that cannot be captured by the latent connectivity matrix \(A\). It should be noted that the inclusion of the last term in Eq. (2.3) distinguishes this model from typical bilinear DCM (see Lemma 1). These matrices altogether provide information about the structural connectivity of a system.
Lemma 1

The second order nonlinear latent connectivity between state variables is given by $\text{diag}(x)Hx$, i.e. the last term in Eq. (2.3).

Proof

Given Eq. (2.7),

$$h_{lk} = \frac{\partial^2 F}{\partial x_l \partial x_k}$$

Denote the column vector: \( \text{diag}(x)Hx = d \)

Then:

$$d_j = \sum_i \left( x_i \delta_{ij} \left( \sum_k h_{lk}x_k \right) \right)$$

$$d_j = \sum_k h_{jk}x_jx_k$$

which is clearly the second order nonlinear latent connectivity between the state variables.

2.3 Parameter Estimation with Bayesian Inferencing.

Given the equations of nonlinearity in state as in Eq. (2.1), it is desired to integrate these to express the outputs directly as a function of inputs and the parameters to be estimated. This has been communicated earlier by (Bankole and Bhattacharyya, 2016) for a special case where the output is a linear combination of inputs. However for the generic nonlinear measurement response model given in Eq. (2.2), model integration proceeds as follows.

2.3.1 Model Integration

Given the equations of nonlinearity in state as in Eq. (2.1), it is desired to integrate these to express the outputs directly as a function of inputs and the parameters to be estimated. This has been communicated earlier by (Bankole and Bhattacharyya, 2016) for a special case where the output is a linear combination of inputs. However for the generic nonlinear measurement response model given in Eq. (2.2), model integration proceeds as follows.
Considering Taylor series expansion of the function \( F(x, u, \theta) \) around a nominal steady state point and collecting terms, Eq. (2.3) can be written as:

\[
\dot{X}(t) = \left( \tilde{A} + \sum_j u_j \tilde{B}_j + \text{diag}(X) \tilde{H} \right) X
\]  

(2.8)

\[
X = \begin{bmatrix} 1 \\ x \end{bmatrix}
\]  

(2.9)

\[
\tilde{A} = \begin{bmatrix} 0 & 0 \\ F(x_0, 0, \theta) & \frac{\partial F(x_0, 0, \theta)}{\partial x} \end{bmatrix}
\]  

(2.10)

\[
\tilde{A} = \begin{bmatrix} 0 & 0 \\ 0 & A \end{bmatrix}
\]  

(2.11)

\[
\tilde{B}_j = \begin{bmatrix} 0 & 0 \\ \frac{\partial F(x_0, 0, \theta)}{\partial u_j} & \frac{\partial^2 F(x_0, 0, \theta)}{\partial x \partial u_j} \end{bmatrix}
\]  

(2.12)

\[
\tilde{B}^j = \begin{bmatrix} \text{col}_j[C] & 0 \\ B_j \end{bmatrix}
\]  

(2.13)

\[
\tilde{H} = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\partial^2 F(x_0, 0, \theta)}{\partial x^2} \end{bmatrix}
\]  

(2.14)

\[
\tilde{H} = \begin{bmatrix} 0 & 0 \\ 0 & H \end{bmatrix}
\]  

(2.15)

The matrix \( \tilde{A} \) is a concatenated form of the matrix \( A \). The matrix \( \tilde{B}_j \) corresponds to the \( j \)th input and includes both the induced connectivity matrix \( B_j \) and the \( j \)th column of the matrix \( C \). \( \text{col}_j[C] \) represents the \( j \)th column of the \( C \) matrix.

Having Eq. (2.8) written in the linear form with \( \tilde{A}, \tilde{B}^j \) and \( \tilde{H} \) defined as in Eqs. (2.10), (2.12) and (2.14) and augmenting the state space vector as in Eq. (2.9), the resulting set of ordinary differential equations (ODEs) can be solved by the customary solution of the first order ODE using the matrix exponential. Assuming that the inputs \( u(t) \) and state space vector \( x \) are relatively constant for a small time interval \( \Delta t \), then the argument of the matrix exponential can be treated as a constant
over the time interval $\Delta t$ therefore this allows a quasi-analytical integration of the ODE given by Eq. (2.8) from $0 \rightarrow \Delta t$ to yield:

$$X(\Delta t) \approx \exp \left( \Delta t \left( \bar{A} + \sum_j \bar{B}_j u_j(0) + \text{diag}(X(0))\bar{H} \right) \right) \cdot X(0) \quad (2.16)$$

Performing the above iteratively and generalizing for any time $t = T(\Delta t)$, Eq. (2.16) becomes:

$$X(t) = X(T\Delta t) \approx \prod_{k=0}^{T-1} \exp \left( \Delta t \left( \bar{A} + \sum_j \bar{B}_j u_j(k\Delta t) + \text{diag}(X(k\Delta t))\bar{H} \right) \right) \cdot X(0) \quad (2.17)$$

Therefore the output responses can be obtained as:

$$Y(t) \approx g \left( \prod_{k=0}^{T-1} \exp \left( \Delta t \left( \bar{A} + \sum_j \bar{B}_j u_j(k\Delta t) + \text{diag}(X(k\Delta t))\bar{H} \right) \right) \cdot X(0) \right) = h(u, \theta) \quad (2.18)$$

### 2.3.2 Bayesian Inferencing

Bayesian estimation or inference is widely used for system identification and parameter estimation. In this framework, priors are defined with a probability distribution to obtain estimates of the unknown parameters as posterior distribution. One typical approach to expressing these priors is through the Gaussian distribution. Thus upon incorporating the priors with the likelihood, it is sought to find the first and second moments of the Gaussian densities of the parameters. The maximum a posteriori (MAP) in the Bayesian framework is considered here.

Consider the estimation of a random parameter vector $\theta$. From Bayes’ rule, the posterior density function (pdf) of the parameter vector $\theta$ given the output response $y$, $\pi(\theta|y)$ is given by:

$$\pi(\theta|y) = \frac{l(y|\theta)p(\theta)}{m(y)} \quad (2.19)$$

Where
The first term \( l(y|\theta) \) in Eq. (2.19) is the likelihood of the parameter vector \( \theta \), while the second term \( p(\theta) \) is the prior probability of the parameter vector. Both of these influence the posterior and the relative influence of each depends on the mode, variance and skewness of their probability density functions. The posterior distribution is also seen as the ratio of the joint distribution (of the output response and the parameter vector) and the marginal distribution of the output vector.

It is desired to obtain optimal estimate of the parameters of the model in Eq. (2.1) and its hyperparameters jointly using the dynamic data with additive measurement noise \( \omega \sim \mathcal{N}(0, C_\omega) \). Various methods have been proposed to solve this type of problem including the principal component analysis of the total least squares (Golub and Van Loan, 1980) and subtraction of the noise statistics in the magnitude spectral domain (Boll, 1979). In sequential approaches, filtering techniques such as the extended Kalman filter or the unscented Kalman filter are used (Wan and Van Der Merwe, 2000). Since the parameters space is stochastic for the given problem, the process is nonlinear, and the process and measurement noises are not necessarily Gaussian, a Bayesian approach is used here. The Bayesian approach also facilitates to cast the user belief in form of priors. Since it is sought to optimally estimate the parameters pertaining to the connectivity matrices that are constrained by the physical configuration of the process, the priors help to realize the physically plausible connectivity parameters. If the user knowledge is available for certain parameters, then those parameters are treated as informed priors. They are assigned a smaller variance while the variance of the uninformed priors are set at a higher value.

In the following equations, the model prediction is given as \( h(u, \theta) \), the measured output is given as \( y \) while the estimate of the underlying signal, i.e. raw output without noise is denoted by \( \hat{y} \). The residual between the model prediction \( h(u, \theta) \) and the estimate of \( y \) is denoted by \( \epsilon \) as shown by Eq. (2.21) with a Gaussian distribution given by \( \epsilon \sim \mathcal{N}(0, C_\epsilon) \).

\[
y = h(u, \theta) + \epsilon
\]

As the measured output \( y \) is only available, the underlying signal \( \hat{y} \) and the corresponding parameters \( \theta \) must be estimated conditioned on the measured output. The joint probability distribution of the state and the parameters given by \( p(\hat{y}, \theta|y) \) can be estimated using Bayes’ law as given below:
\[ p(\hat{y}, \theta | y) = \frac{p(y|\hat{y}, \theta)p(\hat{y}|\theta)p(\theta)}{p(y)} \] (2.22)

Taking logarithm of both sides yields:

\[ \ln(p(\hat{y}, \theta | y)) = \ln(p(y|\hat{y}, \theta)) + \ln(p(\hat{y}|\theta)) + \ln(p(\theta)) - \ln(p(y)) \] (2.23)

Given the above, the underlying signal \( \hat{y} \) and the parameters \( \theta \) that maximize the joint probability \( p(\hat{y}, \theta | y) \) are found such that:

\[ \hat{y}^*, \theta^* = \max_{\hat{y}, \theta} p(\hat{y}, \theta | y) \] (2.24)

To begin, priors are assumed on the parameters given by \( \theta \sim \mathcal{N}(\eta_\theta, C_\theta) \) where \( \eta_\theta \) denotes the prior mean. Thus the formulation results in a maximum a posteriori estimate of the parameter vector and a maximum likelihood estimate of the underlying output signal in the absence of priors for the output signal. The difference between the maximum likelihood and the maximum a posteriori is the presence of priors in the latter. One approach for obtaining optimal estimates of our unknown parameters and the underlying signal is to directly maximize the joint estimation while seeking to obtain the output vector and the parameter estimates in one step, several authors have reported convergence problems with this approach (Nelson and Stear, 1976). In this framework, the signal is estimated given the conditional estimates of the parameters as discussed in subsection 2.3.3 while estimation of the parameters proceeds recursively through a two-step expectation maximization scheme as discussed in Section 2.3.4.

### 2.3.3 Signal Estimation

Suppose the set of all estimate \( \hat{y} \) is given as \( \mathcal{Y} \), then the maximum likelihood estimate of a signal \( y \) is given as \( \hat{y}^* \) where:

\[ \hat{y}^* \in \{ \hat{y} \in \mathcal{Y} : p(\hat{y}|y; \hat{\theta}) \geq p(\hat{\theta}|y; \hat{\theta}) \ \forall \ \hat{y} \in \mathcal{Y} \} \] (2.25)

The maximum likelihood estimate of \( \hat{y}^* \) can be obtained by setting the derivative of the log likelihood with respect to state vector to zero as follows:

\[ \frac{\partial \ln(p(\hat{y}^*|y; \hat{\theta}))}{\partial \hat{y}} = 0 \] (2.26)
\[
\hat{y}^* = (C_y^{-1} + C_e^{-1})^{-1} \left( C_y^{-1} y + C_e^{-1} h(u, \hat{\theta}) \right)
\]  

(2.27)

As can be seen from Eq. (2.27), the expected values of the underlying signal is a weighted estimate of the raw observation and the model prediction weighted by the inverse of the covariance matrices. In addition, as is usually encountered with industrial data, some of the measurements of \( y \) may be missing and these are replaced with \( \mathbb{E}(\hat{y}) \) where \( \mathbb{E}(\cdot) \) denotes expectation.

### 2.3.4 Expectation Maximization

The expectation maximization algorithm (EM) is a generic, iterative algorithm for jointly estimating parameters and hyperparameters of a model (Dempster et al., 1977, Dempster et al., 1981). Originally introduced by Hartley (Hartley, 1958) as an iterative method to obtain maximum likelihood estimates of parameters in the presence of missing data, it was then used in (Orchard and Woodbury, 1972) where theoretical foundation of the underlying idea was provided. It was generalized by (Dempster et al., 1977) where the general results about the behavior of the algorithm as well as a variety of applications were provided. Recently, a formulation of the EM by (Neal and Hinton, 1998) relates the iterative procedure in the EM as a coordinate descent on the free energy of the system.

#### 2.3.4.1 Expectation: Parameter mean and covariance estimation

Upon finding the maximum likelihood estimate of the underlying signal (noise free outputs) from Eq. (2.27), the parameters and hyperparameters of the error covariance matrix are estimated. The hyperparameters are used to parameterize the covariance matrix as shown in Eq. (2.28) At the \( j^{th} \) iteration, let the conditional expectation of the parameters be denoted by \( \eta_{\theta|y} \). The unknown error covariance \( C_e \) is parameterized as follows:

\[
C_e = \sum_k \lambda_k V_k
\]  

(2.28)

The constants \( \lambda_k \) are known as hyperparameters as they scale the contribution of the basis matrices \( V_k \) to the error covariance matrix. The matrix \( V_k \) are sparse matrix with the \( k^{th} \) element in the leading diagonal equal to one and all other elements set to zero. With this parameterization, one can obtain the diagonal covariance matrix underlying the variances. The representation above can also be seen as the basis sets being equal to the first partial derivative of the error covariance matrix.
with respect to the hyperparameters wherein $V_k = \partial C_\epsilon / \partial \lambda_k$. The parameters $\theta$ are estimated by performing a gradient ascent on the joint log likelihood function $\ln(p(\hat{y}, \theta | y))$ to obtain the following recursive identification.

$$
\eta^{j+1}_{\theta | \hat{y}} = \eta^j_{\theta | \hat{y}} - \left( \frac{\partial^2 \ln(p(\hat{y}, \theta | y))}{\partial \theta^2} \right)^{-1} \frac{\partial \ln(p(\hat{y}, \theta | y))}{\partial \theta} \tag{2.29}
$$

By performing a local linear Taylor series approximation at a current estimate of $\theta$, such that

$$
\hat{y} - h(u, \eta^j_{\theta | \hat{y}}) \approx \frac{\partial h(u, \eta^j_{\theta | \hat{y}})}{\partial \theta} \Delta \theta + \epsilon \tag{2.30}
$$

This results in the framework:

$$
J = \frac{\partial h(u, \eta^j_{\theta | \hat{y}})}{\partial \theta} \tag{2.31}
$$

$$
\bar{y} = \begin{bmatrix}
\hat{y} - h(u, \eta^j_{\theta | \hat{y}})
\eta_{\theta} - \eta^j_{\theta | \hat{y}}
\end{bmatrix} \tag{2.32}
$$

$$
\bar{J} = \begin{bmatrix}
J \\
1
\end{bmatrix} \tag{2.33}
$$

$$
C_\epsilon = \begin{bmatrix}
\Sigma \lambda_k V_k & 0 \\
0 & C_\theta
\end{bmatrix} \tag{2.34}
$$

$$
C_{\theta | \bar{y}} = (\bar{J}^T C_\epsilon^{-1} \bar{J})^{-1} \tag{2.35}
$$

$$
\Delta \eta_{\theta | \bar{y}} = C_{\theta | \bar{y}} (\bar{J}^T C_\epsilon^{-1} \bar{y}) \tag{2.36}
$$

$$
\eta^{j+1}_{\theta | \bar{y}} = \eta^j_{\theta | \bar{y}} + \Delta \eta_{\theta | \bar{y}} \tag{2.37}
$$

Eq. (2.31-2.37) reduces to a Gauss-Newton method of nonlinear parameter estimation in the absence of priors, if however the priors are flat and the function is linear, the scheme represents the minimum variance classical Gauss Markov estimator which finds the parameters that minimize the variance or Mahanalobis distance of the data to the model (Friston, 2002).
2.3.4.2 Maximization: Covariance component estimation

Once the parameters are obtained, the hyperparameters employed in the component of the error covariance matrix must be identified. This follows from a maximum likelihood approach which maximizes the likelihood of the conditional estimate of the data obtained from the E step on the current estimate of the hyperparameters. Denoted by \( p(y|\lambda) \), this is obtained by integrating out the dependence of the likelihood on unknown parameters \( \theta \) using the conditional distribution \( q(\theta) \) (Friston et al., 2003). Again, the \( j^{\text{th}} \) iteration of the hyperparameter \( \lambda \) is denoted by \( \eta_{\lambda|y}^j \)

\[
\ln(p(\hat{y}|\lambda)) = \ln \int q(\theta) \frac{p(\hat{y}, \theta|\lambda)}{q(\theta)} d\theta
\]

Using Jensens inequality, the above expression is replaced with a tractable function \( \bar{F} \) given as a lower bound such that:

\[
\ln \int q(\theta) \frac{p(\hat{y}, \theta|\lambda)}{q(\theta)} d\theta \geq \bar{F} = \int q(\theta) \ln \frac{p(y, \theta|\lambda)}{q(\theta)} d\theta
\]

As above, estimation of the hyperparameters proceed from a gradient ascent on the log function \( \bar{F} \) to give the following recursive estimation (Friston et al., 2002, Harville, 1977):

\[
\eta_{\lambda|y}^{j+1} = \eta_{\lambda|y}^j - \alpha \left( \frac{\partial^2 \bar{F}}{\partial \lambda} \right)^{-1} \left( \frac{\partial \bar{F}}{\partial \lambda} \right)
\]

\[
\left( \frac{\partial \bar{F}}{\partial \lambda} \right)_i = \frac{1}{2} \text{trace} \{ MV_j \} - \frac{1}{2} \bar{y}^TM^TV_iM \bar{y}^T
\]

\[
\left( \frac{\partial^2 \bar{F}}{\partial \beta^2} \right)_{ij} = \frac{1}{2} \text{trace} \{ MV_iMV_j \}
\]

where \( M \) is defined as \( C_{\epsilon}^{-1} - C_{\theta|y}^{-1}f^T C_{\epsilon}^{-1} \). Here \( \alpha \) is chosen such that \( 0.5 \leq \alpha < 1 \) to ensure numerical stability.

The results of the expectation maximization algorithm are numerical entries into the A, B, C and H matrix of the second order nonlinear model corresponding to the observed data and any priors enforced into the scheme. The algorithm is terminated based on first order optimality of \( f(\theta) \) or residual between consecutive estimates of \( y \) and/or \( \theta \). While \( a_{ij} \) represents the latent connectivity from \( x_j \) to \( x_i \), \( a_{ii} \) represents latent connectivity of \( x_i \) with itself, therefore connectivity strength is
inferred by comparing off diagonal elements to diagonal elements. This is applicable to elements of matrices $B$ and $H$ as given by Table 2.1 below:

**Table 2.1** Classification of strength for latent and induced connectivity

<table>
<thead>
<tr>
<th>Condition</th>
<th>Latent connectivity(linear)</th>
<th>Latent connectivity(nonlinear)</th>
<th>Induced connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>a_{ij}/a_{ii}</td>
<td>\geq \zeta$</td>
<td>Strong</td>
</tr>
<tr>
<td>$</td>
<td>a_{ij}/a_{ii}</td>
<td>&lt; \zeta$</td>
<td>Weak</td>
</tr>
<tr>
<td>$</td>
<td>h_{ij}/h_{ii}</td>
<td>\geq \zeta$</td>
<td>N/A</td>
</tr>
<tr>
<td>$</td>
<td>h_{ij}/h_{ii}</td>
<td>&lt; \zeta$</td>
<td>N/A</td>
</tr>
<tr>
<td>$</td>
<td>b_{ij}/b_{ii}</td>
<td>\geq \zeta$</td>
<td>N/A</td>
</tr>
<tr>
<td>$</td>
<td>b_{ij}/b_{ii}</td>
<td>&lt; \zeta$</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The connectivity threshold parameter $\zeta$ is set by the user for determining connectivity strength.

### 2.4 Results & Discussions

#### 2.4.1 Toy Example

This toy example is for the purpose of illustration and comparison with existing literature methods to validate the competitiveness of our proposed approach with respect to computational complexity and accuracy. Other examples illustrate the use of the algorithms for connectivity as earlier described. Consider the discrete time model below

$$
\begin{align*}
    x(t + 1) &= ax(t) + b(u(t) \cdot x(t)) + cu(t) + dx(t)^2 + v(t) \\
    y(t) &= x(t) + \omega(t)
\end{align*}
$$

The true parameters are given as: $a = 2.5, b = -3.25, c = 0.1, d = 2.5$ with $v = \omega = \mathcal{N}(5.0 \times 10^{-7})$. The simulation involves a random signal $u(t)$. The model equation was simulated for a 100 points and estimation of parameters was performed using the algorithm presented in this chapter and (Schön et al., 2011) with 10 particles and 100 particles. The results are presented in Fig. 2.2 and Table 2.2.
Figure 2.2 Comparison of parameter estimation computation time.

All algorithms are run on an Intel® Xeon® CPU E-5-1620 v2 with 32GB RAM. While our method converges in 15 iterations in 104s, the algorithm in (Schön et al., 2011) converges in 500 iterations for 10 and 100 particles with runtime of 15hrs and 12hrs respectively. This computational expense is due to the use of particle filtering and smoothing. The computational time is an order of magnitude higher than the algorithm presented here. Our algorithm avoids this additional layer of complexity with superior results for the system defined in Eq. (2.43). This is shown in Table 2.2.

Table 2.2 True and estimated parameter values for different algorithms

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>This thesis</th>
<th>Schön et al (10 particles)</th>
<th>Schön et al (100 particles)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>2.5</td>
<td>2.495001</td>
<td>1.760220</td>
<td>3.068775</td>
</tr>
<tr>
<td>$b$</td>
<td>$-3.25$</td>
<td>-3.131596</td>
<td>-6.720955</td>
<td>-3.286073</td>
</tr>
<tr>
<td>$c$</td>
<td>0.1</td>
<td>0.158269</td>
<td>0.066202</td>
<td>-0.009576</td>
</tr>
<tr>
<td>$d$</td>
<td>2.5</td>
<td>2.491762</td>
<td>2.815121</td>
<td>2.781233</td>
</tr>
</tbody>
</table>
2.4.2 Van De Vusse Reactor with Separator

The Van de Vusse reactor problem is a benchmark problem for nonlinear control case studies in the open literature (Chen et al., 1995, Vojtesek and Dostal, 2010). The reactor is a stirred tank reactor with a cooling jacket which maintains the reactor temperature by removing excess heat produced due to the chemical reactions. The reactor is continuously fed with an input stream with temperature $T_{in}$ containing the reactant cyclopentadiene with a concentration $C_F$. In the liquid phase of the reactor, the reactions consist of a main reaction involving the conversion of cyclopentadiene (species ‘P’) to the product cyclopentenol (species ‘Q’). The main reactant cyclopentadiene also reacts in an unwanted parallel reaction to produce a byproduct dicyclopentadiene (species ‘S’). Additionally, the product cyclopentenol also reacts in an unwanted reaction to form cyclopentanediol (species ‘R’). This reaction scheme is referred to as the Van de Vusse reaction scheme and is described by the set of equations below:

\[
\begin{align*}
  P & \xrightarrow{k_1} Q & \xrightarrow{k_2} R \\
  2P & \xrightarrow{k_3} S
\end{align*}
\]

In addition to the reactor that is typically used in the open literature, a separator is added in this case study as shown in Fig. 2.3.

Figure 2.3 Reactor Separator set up for the Van De Vusse reactor
Unreacted cyclopentadiene (species ‘P”) from the separator is recycled back to the reactor. The separator does not represent any real equipment item, but is introduced so that the developed algorithm can be tested in presence of recycle streams that affects the connectivity of the system, for a process of which the ‘true’ connectivity matrices are known. In this system, the elements of the connectivity matrices here infer the forward and backward connections between the state space variables of the reactor-separator system. Parameters shown in Table 2.3 have been taken from (Bequette, 2003).

\[
\dot{C}_P = F_V (C_{Pf} - C_P) - k_1 C_P - k_3 C_P^2 + R_V (C_{Pr} - C_P) \tag{2.44}
\]

\[
\dot{C}_Q = -F_V C_Q + k_1 C_Q - k_2 C_Q + R_V (C_{Qr} - C_Q) \tag{2.45}
\]

\[
\dot{C}_{Pr} = \frac{V_{rx}}{V_{sep}} (F_V + R_V) (C_P - C_{Pr}) \tag{2.46}
\]

\[
\dot{C}_{Qr} = \frac{V_{rx}}{V_{sep}} (F_V + R_V) (C_Q - C_{Qr}) \tag{2.47}
\]

C denotes concentration while the subscripts denote the species, the additional subscript ‘r’ and ‘f’ denote recycle and feed streams, respectively. \(V_{rx}\) and \(V_{sep}\) are the reactor and separator volumes, respectively. Data used for the algorithm as described are generated by simulating the nonlinear model and subsequently measurements are obtained for the four output variables \(C_P, C_Q, C_{Pr}, C_{Qr}\).

For this particular example, ‘true’ values of the parameters in the connectivity matrices for the second order nonlinear model can be obtained from the full nonlinear model by following the method described in Section 2.3.1. Following equations are obtained:

\[
\dot{x} = Ax + \sum u_j B^j x + \text{diag}(x) H x \tag{2.48}
\]

\[
x = [C_P \ C_Q \ C_{Pr} \ C_{Qr}]^T, \tag{2.49}
\]

\[
u = [F_V \ R_V]^T \tag{2.50}
\]

Informed priors are assigned a smaller variance with a value of 0.1 while the variance of the uninformed priors are set at 10.0.
Table 2.3 Parameters used for the Van de Vusse reactor separator

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_V )</td>
<td>Recycle flow rate to reactor volume ratio</td>
<td>( C_P )</td>
<td>Concentration of species P</td>
</tr>
<tr>
<td>( \frac{V_{rx}}{V_{sep}} )</td>
<td>Ratio of reactor volume to separator volume</td>
<td>( C_{Pr} )</td>
<td>Concentration of species P in recycle stream</td>
</tr>
<tr>
<td>( k_1 )</td>
<td>Reaction 1 rate constant</td>
<td>( C_Q )</td>
<td>Concentration of species Q</td>
</tr>
<tr>
<td>( k_2 )</td>
<td>Reaction 2 rate constant</td>
<td>( C_{Qr} )</td>
<td>Concentration of species Q in recycle stream</td>
</tr>
<tr>
<td>( k_3 )</td>
<td>Reaction 3 rate constant</td>
<td>( F_V )</td>
<td>Feed flow rate to reactor volume ratio</td>
</tr>
</tbody>
</table>

For identification purposes, the nonlinear model is simulated with additive Gaussian noise. White noise was added to the raw data obtained from the simulation of the equation depicting the true dynamics of the system.

![Figure 2.4 Concentration profile of P (left), Q (right) for a signal to noise ratio of 24](image)

Figure 2.4 Concentration profile of P (left), Q (right) for a signal to noise ratio of 24
Figure 2.5 Euclidean distance between noisy signal and signal estimate (left), Euclidean distance between true parameter vector and estimates (right), signal to noise ratio of 24.

Here, the A, C, H matrices consist of 16, 8, 16 elements, respectively while each of the B matrices consist of 16 elements totaling 72 elements. Fig. 2.4 shows that the estimates of the profiles of concentration of species match the true data very well even in the presence of noise. For brevity, the concentration profiles of unreacted cyclopentadiene (Cₚᵣ) and recycled cyclopentenol (Cₚᵣ) are omitted.

Fig. 2.5 shows that the Euclidean distance between the observed data and the underlying state as well as the 2-norm difference between the true parameters and the estimated parameters are seen to decrease with iterations. However for a lower level of noise in the data with a signal to noise ratio of 30 (not shown), eight (8) iterations are needed for the algorithm to converge.

Lastly, with the connectivity threshold parameter \( \zeta = 0.8 \) set, the strength of coupling between species in the reactor and separator are evaluated. It is observed that strong latent connectivity exists between the reactor and the separator due to species \( C_p \) and \( C_p'r \). Also strong induced connectivity exists between the reactor and separator between the species \( C_p \) and \( C_p'r \) both due to the feed flow rate \( F_v \) and the recycle flowrate \( R_v \), these are based upon the numerical estimates of matrices \( B_1 \) (due to \( F_v \)) and \( B_2 \) (due to \( R_v \)) , respectively. However in comparison to the true structural connectivity, weak latent connectivity is inferred from the reactor to separator between species \( C_Q \) and \( C_Q'r \) and weak induced connectivity is inferred from the reactor to separator between species \( C_Q \) and \( C_Q'r \) due to changes in the feed flow rate.
Table 2.4 Latent and Induced connectivity

<table>
<thead>
<tr>
<th>Species</th>
<th>Parameter</th>
<th>Estimate</th>
<th>True</th>
<th>Parameter</th>
<th>Estimate</th>
<th>True</th>
<th>Estimate</th>
<th>True</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_p, C_{pr}$</td>
<td>$</td>
<td>a_{31}/a_{33}</td>
<td>$</td>
<td>1.05</td>
<td>1.00</td>
<td>$</td>
<td>b_{31}/b_{33}</td>
<td>$</td>
</tr>
<tr>
<td>$C_Q, C_{qr}$</td>
<td>$</td>
<td>a_{42}/a_{44}</td>
<td>$</td>
<td>0.97</td>
<td>1.00</td>
<td>$</td>
<td>b_{42}/b_{44}</td>
<td>$</td>
</tr>
</tbody>
</table>

2.4.3 Acid Gas Removal Unit

The integrated gasification and combined cycle (IGCC) unit (Bhattacharyya et al., 2010) is evaluated. This system presents a good study for examining the dynamic causal model due to strong mass/heat interactions and high nonlinearities. The gasifier produces syngas, mainly hydrogen and carbon monoxide, which is sent to a series of water gas shift reactors (modeled as adiabatic plug flow reactors in series) with inter-stage cooling. The shifted syngas is then sent to the acid gas removal unit where acid gases (CO$_2$ and H$_2$S) are absorbed from the dirty syngas leaving mainly hydrogen in the clean syngas. The cleaned syngas is then sent to the gas turbine for power production. The dynamic causal model is implemented on the acid gas removal (AGR) unit. A detailed analysis of the process can be found in (Bhattacharyya et al., 2010).

![Process flow configuration of the acid gas removal unit](image-url)
The acid gas removal unit model used here is available in DYNSIM® (http://software.schneider-electric.com/) and is divided into three sections, each section being run in a separate DYNSIM engine and each DYNSIM engine is run on a different processor enabling distributed computing.

The development of the dynamic causal model for the AGR unit proceeds as follows: first, all the unit operations (excluding utilities) in the process flowsheet in each engine are identified. Secondly, pertinent variables used for the characterization of the system are identified. In the AGR unit, these include the vapor and liquid composition of the streams in terms of the major species of the system i.e. hydrogen ($H_2$), carbon dioxide ($CO_2$) and hydrogen sulfide ($H_2S$) as well as temperature ($T$) of these streams. For each unit operation, flow rates of all incoming flow streams are designated as extrinsic inputs while other variables such as species concentration and temperatures are denoted as induced variables only if the streams corresponding to those variables emanate from another unit operation. It should however be noted that temperature is not considered as a candidate variable for induced connectivity between unit operations with a heat exchanger between them but rather as an extrinsic input. For illustration, Table 2.2 summarizes the classification of inputs as extrinsic and induced as well as their sources for all unit operations in the CO$_2$ absorber flowsheet only. Due to the restriction imposed on the connectivity matrix by the physical configuration of the process setup, a number of priors are enforced into the scheme. This prevents the realization of physically implausible connectivity parameters. As an illustration, unconnected unit operations have all latent, induced and extrinsic connectivity elements set to zero.

The model was run from steady state with perturbation in the incoming CO$_2$-laden syngas flowrate and with 20% deviation from steady state. For this simulation, 1000 data points (between $t = 0$ and $t = 250$ s) were collected from the simulation and the output variables were scaled with respect to the maximum value (Appendix A.2). The normalized data were then used in the Bayesian framework as outlined in subsection 2.3.2. For simplicity, only a few results in Figs. 2.7 and 2.8 will be shown to indicate how the estimated profiles compare with the raw data.
Table 2.5 Classification of inputs and state variables for unit operations in CO$_2$ absorber flowsheet

<table>
<thead>
<tr>
<th>Tag</th>
<th>Equipment</th>
<th>Extrinsic input</th>
<th>Induced input</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>CO$_2$ absorber</td>
<td>Semilean solvent flowrate</td>
<td>Semilean solvent species concentration</td>
<td>Low pressure flash drum (D4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lean solvent flowrate</td>
<td>Lean solvent species concentration</td>
<td>Selexol stripper (T3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hydrogen recovery flowrate</td>
<td>Hydrogen recovery species concentration</td>
<td>Hydrogen recovery knock out drum (D1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vapor flowrate from H$_2$S absorber</td>
<td>H$_2$S absorber top tray vapor species concentration</td>
<td>H$_2$S absorber (T2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1</td>
<td>Hydrogen recovery knock out drum</td>
<td>Hydrogen flowrate from recovery drum</td>
<td>Hydrogen recovery drum species concentration</td>
<td>Hydrogen recovery drum (D2)</td>
</tr>
<tr>
<td>D2</td>
<td>Hydrogen recovery drum</td>
<td>Liquid flowrate from CO$_2$ absorber</td>
<td>CO$_2$ absorber liquid phase species concentration</td>
<td>CO$_2$ absorber (T1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D3</td>
<td>MP CO$_2$ flash drum</td>
<td>Liquid flowrate from Hydrogen recovery drum</td>
<td>Hydrogen recovery liquid phase species concentration</td>
<td>Hydrogen recovery drum (D2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D4</td>
<td>LP CO$_2$ flash drum</td>
<td>Liquid flowrate from MP CO$_2$ flash drum</td>
<td>MP flash drum liquid phase species concentration</td>
<td>MP CO$_2$ flash drum (D3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temperature</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.7 Normalized vapor phase concentration of CO$_2$ in stripper (left), liquid phase concentration of H$_2$S (right) in the CO$_2$ absorber: true data (blue dash dot), estimates (solid red), noisy signal (black).

Figure 2.8 Normalized LP flash drum temperature (left) and H$_2$S concentrator sump temperature (right).

The algorithm presented in Section 2.3.4 is run on an Intel® Xeon® CPU E-5-1620 v2 with 32GB RAM which took 45 minutes. The obtained induced connectivity results are summarized in Table 2.6, latent connectivity is shown in Table A.1. As in the case of the Van De Vusse reactor, a connectivity threshold of $\zeta = 0.8$ is considered. A few observations can be made in Table 2.6. With respect to the CO$_2$ absorber (T1) row, strong connectivity is observed from H$_2$ recovery...
knock out drum to the CO₂ absorber with respect to all variables except H₂ in liquid phase and H₂S. The weak connectivity due to H₂S can be attributed to the absorption in the H₂S absorber as most of the H₂S in the incoming syngas stream (≥ 95%) is absorbed by the incoming solvent from the CO₂ absorber. This is expected since the process is designed for deep removal of H₂S (Bhattacharyya et al., 2010). The weak connectivity due to H₂ in the liquid phase can be explained by noting that the H₂ in the incoming vapor phase stream form the hydrogen recovery drum is fed towards the top of the tower thus having a lesser effect. However with respect to the low pressure flash vessel, strong connectivity is observed with the CO₂ absorber, this is due to the fact that the operating conditions of the low pressure flash drum significantly affects CO₂ capture in the absorber as discussed in (Bhattacharyya et al., 2010).

Weak connectivity is observed from the H₂S absorber to the CO₂ absorber for H₂S since most of the H₂S gets captured in the H₂S absorber. On the other hand, a strong connectivity is observed from the H₂S absorber to the CO₂ absorber for CO₂ capture since most of the CO₂ capture does take place in the CO₂ absorber. On the other hand, strong connectivity from the CO₂ absorber to the H₂S absorber is observed as expected since the species concentration of the solvent leaving the CO₂ absorber greatly affects that of the H₂S absorber. The medium pressure (MP) flash drum (D3) is weakly connected to the H₂ recovery drum while it is only strongly connected to the low pressure (LP) flash drum due to temperature. As flow progresses downstream from the CO₂ absorber to the H₂S stripper, a decrease in the connectivity due to CO₂ is observed, which is expected since most of the CO₂ is captured in the CO₂ absorber and then stripped off from the solvent in the flash vessels in the flash vessels (D2-D4). For this test case, only a qualitative comparison could be made. Connectivity information obtained from the EM algorithm is found to be at par with the underlying thermodynamic and first-principles model. While this threshold ζ is chosen based on heuristic, several runs were performed on the results by varying the connectivity threshold as shown in Fig. 2.4, as expected an increase in the threshold parameter results in a lower proportion of strongly connected variables. This will result in more islands or group of variables and a greater decomposition but could sacrifice the accuracy of representation. Conversely, a lower value of ζ would result in a reduced number of independent islands/groups. This would result in a more accurate depiction of the system but reduced gain for a divide and conquer based structural decomposition.
Table 2.6 Summary of induced connectivity table for the AGR unit*

<table>
<thead>
<tr>
<th>TO</th>
<th>FROM</th>
<th>Variables</th>
<th>CO₂ absorber (T1)</th>
<th>H₂ recovery KO drum (D1)</th>
<th>H₂ recovery drum (D2)</th>
<th>MP flash (D3)</th>
<th>LP flash (D4)</th>
<th>H₂S absorber (T2)</th>
<th>H₂S conc. (T4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂ Absorber (T1)</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂S</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H₂ rec KO drum (D2)</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MP flash (D3)</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP flash(D4)</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H₂S Absorber (T2)</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H₂S concentrator (T4)</td>
<td>CO₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>H₂</td>
<td>Vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Green band denotes strong connectivity while grey band denotes weak connectivity.
2.4.4 Conclusions

It is observed that results from this deployment on the Van de Vusse reactor with separator shows that both strong and weak connectivity are correctly identified within specified tolerances. For the acid gas removal unit, the estimated profiles are in agreement with the true underlying data and the structural connectivity results are found to be qualitatively satisfactory. It should be noted that a different structural connectivity table can be obtained by varying the connectivity threshold parameter $\zeta$. For a very low threshold parameter, it is expected that all units would be considered simultaneously making the large-scale problem computationally intractable. Since solving the decomposed problem might lead to a sub-optimal solution in comparison to when the entire system is solved simultaneously, a high threshold parameter may lead to larger deviation from the optimal solution. Thus the reduction in the computational expense versus the deviation from the optimal solution needs to be weighed. Nevertheless, the method proposed here can be useful where connectivity can be leveraged as a tool for the systematic division of the process into multiple islands making the system amenable to distributed computing and online deployment.
Chapter 3

3 Development of Multi-agent Optimization Based Approach for Controlled Variable Selection

3.1 Background

The previous chapter outlines the algorithm for the division and decomposition of a process into sections/islands. In this chapter, the decomposition will be used as a means to subdivide the process and deploy biologically inspired controlled variable selection on each island in parallel and the results of different islanding/partitioning will be explored in terms of optimality. Control Structure Design has been studied in recent literature with focus on a holistic plant wide approach. This work explores the deployment of controlled variable algorithm for the selection of the optimal set of primary controlled variables on multiple sections of a process plant arising from structural decomposition algorithm. This allows for speedy execution and prospects for faster/online controlled variable selection. Secondly, a metaheuristic based multiagent algorithm is examined as an alternative to traditional branch and bound algorithms for parallelization and improvement in computational speed. Both of these novelties are original works of the author and distinguish this work from (Jones et al., 2014). This platform is employed to solve the mixed integer multi objective optimization selecting controlled variables with promising economic and controllability performance. From each island/section, the results of the controlled variable selection algorithm are merged to form a selection for the whole process. These algorithms and methods are applied to an acid gas removal unit of an integrated gasification combined cycle.

3.2 Introduction

As explained, a combination of two approaches is proposed to reduce the computational time of the second stage optimization problem. The first approach seeks to reduce the size of the optimization problem by reducing the number of combinatorial problems that need to be evaluated. Both the first and second approaches help to reduce the computational time for solving the optimization problem. To proceed with the motivation behind the proposed methodologies, note that the CV selection problem has so far been solved holistically, i.e. the entire plant is evaluated together for CV selection. If the plant can be decomposed into multiple sections, it can result in
significant reduction in the number of possible combinations (a more formal proof is provided later) and the CV selection problem is then solved for each section separately. The decomposition is based on a novel algorithm that partitions the process plant into a number of sections based on the structural connectivity. It can be shown (a more formal proof is provided later) that for a completely unconnected system, the CV selection problem for the original problem collapses to the CV selection problem of the decomposed system. The decomposition provides computational advantage not only due to the reduced number of optimization problems to be solved, but also because the optimization problems can be solved in parallel on multiple processors without any communication overhead among the processors. In the second approach, a multiagent platform is leveraged providing significant computational advantage over the traditionally used BB algorithm. The multiagent platform employs multiple heuristic algorithms facilitating use of homogeneous or heterogeneous agents as needed (Gebreslassie and Diwekar, 2015, Gebreslassie and Diwekar, 2018). The algorithm can also select the optimal agent at any stage of iteration providing further computational advantage.

3.3 Approach

3.3.1 A Priori Analysis

To begin, an objective function $\bar{f}$ must be determined based on the operational objective of the process. The objective function is typically a cost function, profit function, or a measure of plant efficiency that is desired to be optimized. Next, constraints (mainly operational and regulatory), manipulated variables (degrees of freedom) and disturbances are identified. An optimization is performed with respect to the identified degrees of freedom and due consideration of the constraints. These optimizations are undertaken under nominal conditions as well as under various disturbance conditions. These optimization studies yield a number of important information. First, information about the optimal variation of the input and output variables is obtained. This information is used to construct scaling matrices for outputs and inputs, given by Eqs. (3.1) and (3.2), respectively. These scaling matrices are used for maximum singular value rule (Skogestad and Postlethwaite, 2007).

$$
S_y = \text{diag}\left(\max(|c^{\text{nom}} - c^d|)\right) \quad (3.1)
$$

$$
S_u = \text{diag}\left(\max(|u^{\text{nom}} - u^d|)\right) \quad (3.2)
$$
The superscript ‘nom’ and ‘d’ denote nominal and disturbance conditions, respectively, ‘c’ represents controlled variables and ‘u’ represents manipulated variables (MVs). These scaling matrices are used in the next stage where the optimization is formulated.

Second, these optimization studies yield information about the active constraints. These constraints are active in all optimization studies while considering nominal and disturbance conditions. These active constraints must be selected as CVs. Suitable MVs are selected from the available list so that these CVs can be maintained within tight bounds. If there are additional MVs that can be used as degrees of freedom, then additional CVs are selected.

For selecting the additional CVs, first, a list of remaining candidate controlled variables is generated. From this list, prescreening criteria are used to eliminate some of variables to eliminate infeasible CVs thus reducing the size of the combinatorial optimization in the next step. These prescreening criteria can be user dependent. Generally, it would be desired to eliminate variables that exhibit weak servo performance and/or are strongly affected by disturbances. Let $N_y, N_u, N_d$ represent dimensions of measurements, manipulated variables and disturbances respectively. For applying these criteria, a linear process model is obtained from the process under nominal conditions as shown in Eq. (3.3) with $G^y \in \mathbb{R}^{N_y \times N_u}$ as the process gain matrix and $G^y_d \in \mathbb{R}^{N_y \times N_d}$ as the disturbance gain matrix. These gain matrices are scaled such that all elements of inputs $u$, outputs $y$ and disturbances $d$ have a maximum magnitude of 1. The prescreening criteria are mathematically stated in Eqs. (3.4-3.6). If the inequality in Eq. (3.5) is not satisfied, no input can control output variable $y_j$ within the bounds. In addition, candidate controlled variables $y_j$ that have high dead time, represented by $t_d(u_i, y_j)$-beyond a threshold $\chi_j$- with respect to the available manipulated variables $u_i$ can also be prescreened off by using Eq. (3.6). The criterion $\chi_j$ is selected by ordering the time delays estimated in the transfer function model and this is empirically chosen. The pre-screening step can reduce the initial list of candidate controlled variables significantly thus decreasing the size of the combinatorial optimization in the following step.

\[
\begin{align*}
y &= G^y u + G^y_d d \\
\| (G^y)_i \|_\infty &= 1 \ \forall \ i \\
\| (G^y)_i \|_\infty &\geq \| (G^y_d)_i \|_\infty \ \forall \ i \\
t_d(u_i, y_j) &\leq \chi_j
\end{align*}
\]
3.3.2 Formulation of Loss Objective Function (Local Linear Exact Method)

This section presents the derivation of the worst case and the average case loss function which would be evaluated to determine the optimal subset of candidate variables ‘c’ to be chosen as controlled variables in the presence of changing disturbances ‘d’. Here $d \in \mathbb{R}^{N_d}$ represents exogenous and uncontrolled inputs to the system. Given the scalar cost function denoted by $\bar{J}$ (from the first stage of the top down analysis). This scalar cost function is to be minimized by the available degrees of freedom at steady state denoted by $\bar{u}$. Thus the following minimization problem is presented

$$\min_{\bar{u}} \bar{J}(x, \bar{u}, d)$$

Subject to

$$g(x, \bar{u}, d) \leq 0$$
$$h(x, \bar{u}, d) = 0$$

(3.7)

The variable $x$ denotes the states of the system. The degrees of freedom of the overall system is denoted by $\bar{u}$. The constraints include the model equations for the system such as differential algebraic equations $h(x, \bar{u}, d) = 0$ and physical constraints $g(x, \bar{u}, d) \leq 0$. The solution of the optimization problem above results in the separation of active constraints where $g(x, u_{ac}, d) = 0$ and inactive constraints for which $g(x, u, d) < 0$.

It is assumed that the original degrees of freedom can be partitioned as follows, $\bar{u} = \{u_{ac}, u\}$ where $u_{ac}$ consists of the degrees of freedom used for the control of active constraints and $u$ represents the unconstrained degrees of freedom for the unconstrained portion of the optimization as shown in Eq. (3.8). As active constraints are of higher priority, selected manipulated variables are paired with these controlled variables.

$$\min_{u} J(x, u, d)$$

(3.8)

It should be noted that the optimization of $J$ is carried out with respect to unconstrained degrees of freedom as opposed to $\bar{J}$. As the degrees of freedom $u$ may be adjusted to meet the optimal cost function, measurements are made to estimate the disturbances and $u$ is freely adjusted such that the optimal value $u = u_{opt}(d)$ is implemented (as in EMPC). However this approach is nontrivial and requires both the updated value of the disturbances (which may be difficult to measure) as
well as the optimal value of $u$ dependent on $d$. An alternative approach is to keep the controlled variables at constant set point as stated earlier. Given the measured variables $y \in \mathbb{R}^{Ny}$, from this list, a subset $c$ is selected such that $u = u(c, d)$ exists. Given the loss as shown below

$$L(c, d) = J(c, d) - J_{opt}(d)$$  \hspace{1cm} (3.9)

While $J(c, d)$ represents the value of the objective function while keeping variables $c$ controlled, $J_{opt}(d)$ represents the optimal value of $J$. To obtain an expression for the loss function, a local linear analysis is performed (Halvorsen et al., 2003). This is valid for small deviations from the nominal steady state point of operation (denoted by the * superscript). Performing a Taylor series expansion around the nominally operating point, the scalar objective cost function is expressed as:

$$J(u, d) = J^* + J_u^* \Delta u + J_d^* \Delta d + \frac{1}{2} (\Delta u^T J_{uu}^* \Delta u + \Delta d^T J_{dd}^* \Delta d + \Delta d^T J_{ud}^* \Delta u) + \cdots$$  \hspace{1cm} (3.10)

Rewriting in vector form and replacing the deltas with deviation variables, one obtains:

$$J(u, d) = J^* + [J_u^* \text{ } J_d^*][u \text{ } d]^T + \frac{1}{2} \begin{bmatrix} u^T \\ d^T \\ \end{bmatrix} \begin{bmatrix} J_{uu}^* & J_{ud}^* \\ J_{ud}^* & J_{dd}^* \\ \end{bmatrix} \begin{bmatrix} u \\ d \\ \end{bmatrix}$$  \hspace{1cm} (3.11)

At optimality: $J_u^* = J_d^* = 0$. Therefore the difference between the optimal cost function $J^*$ and the true value $J(u, d)$, denoted as loss $L$ is given as:

$$L(d) = J(u, d) - J^* = \frac{1}{2} \begin{bmatrix} u^T \\ d^T \\ \end{bmatrix} \begin{bmatrix} J_{uu}^* & J_{ud}^* \\ J_{ud}^* & J_{dd}^* \\ \end{bmatrix} \begin{bmatrix} u \\ d \\ \end{bmatrix}$$  \hspace{1cm} (3.12)

For optimality of $J$, $u$ must be adjusted whenever disturbances change so that $u_{opt} = u_{opt}(d)$. To arrive at a relationship between the optimal input and the moving disturbance. A local linear model is obtained as shown below in Eq. (3.13). The expression for $u_{opt}$ can be obtained by expanding the first order derivative of the cost function with respect to $u$ around the nominally optimal point.

$$J_u = J_u^* + J_{uu}^* u + J_{ud}^* d$$  \hspace{1cm} (3.13)

As the new point is also optimal, this implies $J_u = J_u^* = 0$, therefore:

$$u_{opt} = -J_{uu}^{-1}J_{ud}^* d$$  \hspace{1cm} (3.14)

Given the model equation $y = f(u, d)$, the linear form is given as:

$$y = G^y u + G_d^y d + n^y$$  \hspace{1cm} (3.15)
Where $G^y = (\partial f / \partial u)$ and $G_d^y = (\partial f / \partial d)$. Therefore $G^y$ represents the gain matrices of the full space of outputs $y$ with respect to the inputs $u$ and $G_d^y$ represents the gain matrices of the full space of outputs $y$ with respect to the disturbances $d$ i.e. $G^y \in \mathbb{R}^{N_y \times N_u}, G_d^y \in \mathbb{R}^{N_y \times N_d}$. $n^y$ is the noise levels of the measured variables $y$. A subset of measured variables chosen as controlled variables $c$ is expressed as:

$$c = H y$$

$$c = Gu + G_d d + n$$

From Eqs. (3.15-3.17), it follows that $G = HG^y$, $G_d = HG_d^y$, $n = Hn^y$. Where the matrix $H$ is the $n_c \times n_y$ matrix mapping from $\mathbb{R}^{n_y} \mapsto \mathbb{R}^{n_c}$ with the condition that rank($HG^y$) = $N_u$. For single measurements, $HH^T = I_{N_u}$. The set points of the optimal controlled variables is denoted by $c_s$ ($c_s = 0$) while the actual measurements is denoted by $c$. Therefore:

$$c_s - c_{opt}(d) = (GJ_u^{-1} J_{ud}^* - G_d) d + n$$

This expression is the difference between the value of the controlled variables at the optimal point and the nominal operating point, therefore the associated change in manipulated variable that is the required driving force of the input to make this correction given in the set point error is given as

$$u_{opt} = G^{-1} \left( (GJ_u^{-1} J_{ud}^* - G_d) d + n \right)$$

Thus at every disturbance $d$, there exists a difference between the optimal input required to keep the controlled variables at their optimal set points and the actual input used to steer the controlled variables to the constant set point obtained for the nominal point. The loss function can be thus expressed as a function of this deviation as shown below:

$$L = J(u(d), d) - J(u^*(d), d)$$

Expanding this in form of Taylor series gives

$$L(d) = \frac{1}{2} u^T J_{uu} u$$

The deviation in input is obtained from (3.19) and substituted into (3.21), the following is obtained:
\[ L = \frac{1}{2} z^T z \]  

(3.22)

where

\[ z = J_{uu}^{1/2} [ (J_{uu}^{-1} J_{ud} - G^{-1} G_d) d - G^{-1} n ] \]  

(3.23)

Scaling the random variables \( d \) and \( n \) with diagonal matrices so that the relative magnitudes of these variables are less than 1, the above expression can be rewritten as:

\[ z = J_{uu}^{1/2} [ (J_{uu}^{-1} J_{ud} - G^{-1} G_d) W_d - G^{-1} W_n ] [d_n] \]  

(3.24)

### 3.3.2.1 Worst and Average Case Loss

#### 3.3.2.1.1 Worst Case Loss

The worst case loss is obtained when the combined value of the disturbances and the measurement noise is 2-norm bounded, which implies the following

\[ \min_{[d_n] \leq 1} L \]  

(3.25)

Defining the worst case loss (Halvorsen et al., 2003) gives:

\[ \sigma(\bar{M})^2/2 \]  

(3.26)

Where

\[ \bar{M} = [M_d M_n] \]  

(3.27)

\[ M_d = J_{uu}^{1/2} (J_{uu}^{-1} J_{ud} - G^{-1} G_d) W_d \]  

(3.28)

\[ M_n = J_{uu}^{1/2} (G^{-1} W_n) \]  

(3.29)

Where, as discussed earlier

\[ G = HG^y; G_d = HG_d^y; W_n = HW_n^y \]  

(3.30)
3.3.2.2 Average Case Loss

The worst case loss as described above can be written as

$$\max_{d \in \mathcal{D}} \max_{n \in \mathcal{N}} \left( J(c, d) - J_{\text{opt}}(d) \right)$$ (3.31)

The worst case may however be an overestimation of the loss case and the average loss over the feasible domain of the disturbances and noise (Kariwala et al., 2008) can be rewritten as:

$$E(L) = \frac{1}{|\mathcal{D}|} \frac{1}{|\mathcal{N}|} \int_{\mathcal{N}} \int_{\mathcal{D}} \left( J(c, d) - J_{\text{opt}}(d) \right)$$ (3.32)

Which gives:

$$E(L) = \frac{1}{2} \left\| M_{d} M_{n} \right\| \frac{d}{n} \left\| \right\|^{2}$$ (3.33)

Where $M_{d}$ and $M_{n}$ are as defined in Eq. (3.28) and Eq. (3.29) above. Given that the deviation of the disturbances belong to a space where the assumed linear model is valid, and that the measurement noise belongs to a set of allowable measurement noises i.e. $d \in \mathcal{D}, n \in \mathcal{N}$. The average loss can therefore be computed as:

$$E(L) = E \left[ \frac{1}{2} \left( \text{tr}(M_{d} M_{n} \bar{d} \bar{d}^{T}) \right) \right]$$ (3.34)

Where $\bar{d} = [d \ n]^{T}$, from the above, one obtains:

$$E(L) = \frac{1}{2} E \left[ \text{tr}(M_{d} \bar{d} \bar{d}^{T}) \right]$$ (3.35)

$$E(L) = \frac{1}{2} M^{T} ME \left[ \bar{d} \bar{d}^{T} \right]$$ (3.36)

Now assuming $\alpha$ is a uniform random variable such that $\alpha = \| \bar{d} \|_{2}, \alpha \sim [0,1]$. $\bar{d} \in \mathbb{R}^{N_{u}+N_{d}}$ .

$$(N_{u} + N_{d}) E \left[ \bar{d} \bar{d}^{T} \right] = \int_{0}^{1} \alpha^{2} \ d\alpha = \frac{1}{3}$$ (3.37)

Therefore substituting Eq. 3.37 for $E \left[ \bar{d} \bar{d}^{T} \right]$ and Eqs. (3.27-3.29) for $M, M_{d}, M_{n}$ respectively in Eq. (3.36), the following is obtained:
\[ L = \frac{1}{6(N_u + N_d)} \left\| J_{uu}^{-1/2} (J_{uu}^{-1} J_{ud} - G_d^{-1} G_d) W_d J_{uu}^{-1/2} G_d^{-1} W_n \right\|_F^2 \]  

(3.38)

Where the subscript \((\cdot)_F\) denotes the Frobenius norm. The difference between this loss function and the average loss presented in (Kariwala et al., 2008) is the scaling factor \(1/(N_u + N_d)\). The average loss function is a measure of the average of the singular values in the matrix hence the division of the Frobenius norm by the rank of the matrix since the number of singular values of any matrix is equivalent to the rank of the matrix. It is noteworthy that the only difference between the average case loss and the worst case loss is the type of norm applied, they are equivalent in every other respect.

### 3.3.3 Controllability

Given the economic measure of the loss function derived above for both worst case loss and average case loss. It is also necessary to quantify the behavior and the optimality of chosen candidate controlled variables in terms of control performance (ease of control). The measure of controllability can be chosen as the inverse of the minimum singular value of the appropriately scaled gain matrix as defined in (Skogestad and Postlethwaite, 2007). This scaled gain matrix is given as

\[ \hat{G} = D_e^{-1} G D_u \]  

(3.39)

Where the diagonal matrices \(D_e\) and \(D_u\) are given by the following expression:

\[ D_e = \text{diag} \left( \min(|c - c^{\text{nom}}|, |c^{\text{nom}} - c|) \right) \]  

(3.40)

\[ D_u = \text{diag} \left( \min(|u - u^{\text{nom}}|, |u^{\text{nom}} - u|) \right) \]  

(3.41)

The superscript ‘nom’ denotes nominal conditions. The singular value decomposition of a matrix \(\hat{G}\) which is an \(N_y \times N_u\) matrix is given by:

\[ \hat{G} = U \Sigma V^T \]  

(3.42)

With \(U\) and \(V\) are \(N_y \times N_u\) and \(N_u \times N_u\) orthogonal matrices. These matrices characterize \(\hat{G}\) such that the columns of \(U\) span the column space of \(G\) and the columns of \(V\) span the row space of matrix \(\hat{G}\). The diagonal matrix \(\Sigma\) is an \(N_y \times N_u\) matrix of rank (\(\min(N_y, N_u)\)) = \(r\) with entries \(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0\). The right singular vectors \(v_j\) which are the columns of \(V\) represent the principal components directions of \(G\) and have the relationship:
\[ \hat{G}v_j = \sigma_j u_j \]  

Thus the diagonal entries represent the ‘gain’ of the system defined for the gain matrix \( \hat{G} \) in a multiple input multiple output (MIMO) sense. Therefore the controllability (ease of control) is defined as a measure of the worst direction of the system under control and taken as the inverse of the smallest singular value of \( \hat{G} \). This is defined as 

\[ J_c(c) = \sigma^{-1}(\hat{G}) \]  

This represents the inverse of the worst input to output gain. Therefore controlled variables ‘c’ should be chosen so as to minimize the quantity \( J_c(c) \).

Other considerations of the controlled variable selection optimization includes the imposition of a user defined constraint which restricts the time delay \( t_d(u_i, y_j) \) between the manipulated variable \( u_i \) and candidate controlled variable \( y_j \) to a certain maximum \( \chi_j \). Additionally, one may impose a subset selection constraint such that only a fixed number of controlled variables may be selected from a subset denoted by columns of \( \Pi \), where \( \Pi \) is a matrix of \( M \) logical vectors of size \( N_y \times 1 \), with 1s for membership. Let \( F_i \) denote the binary variable that candidate \( i \) is picked from the subset and let \( \xi_m \) denote the number of controlled variables that may be picked from subset \( Q_m \). It follows that \( \sum_{i \in \Pi_m} F_i = \xi_m \forall m \in 1, \ldots, M \). The argument that minimizes the combined loss is a logical vector \( P \) denoting the candidate controlled variables chosen. Given the economic and the controllability function, the optimization problem is formulated as follows:

\[
\min_P \{ L, J_c \} \\
\text{Subject to} \\
L = \frac{1}{6(N_u + N_d)} \left\| J^{1/2}_{uu} (J^{-1}_{uu} J_{ud} - G^{-1} G_d) W_d J^{1/2}_{uu} G^{-1} W_n \right\|_F^2 \sum_{i \in \Pi_m} F_i = \xi_m \forall m \in 1, \ldots, M \\
t_d(u_i, y_j) \leq \chi_j
\]  

The development of the multiagent optimization based approach for selection of multiple controlled variable sets is requires some modification as the optimization scheme only returns one global optimum, it is desired to reformulate the problem to find other optimal sets other than the
global in order of decreasing optimality. Thus the problem is reformulated as follows with the pseudocode below:

Pseudocode for generation of unique solutions in multiagent optimization.

Initialize the agents

Initialize an zero matrix \( V \) of size \( N \times P \) where \( N \) is the size of the decision variables and \( P \) is the number of solutions desired

For \( i = 1 \rightarrow P \)

Run the optimization with the cost function

Obtain solution \( x_i = \{x_1, x_2, ..., x_N\}^T \)

For \( j = 1 \rightarrow i - 1 \)

Define \( g_j(x) = \varepsilon - (x_i^T x_i - x_j^T x_j)^2 \)

In the above pseudocode, the variable \( \varepsilon \) satisfies \( 0 < \varepsilon \leq 1 \).

3.3.4 Posteriori Analysis

In addition to the prescreening in the apriori stage and the subsequent economic evaluation of the alternative controlled variable sets, it is pertinent to examine them at off design conditions to screen off candidate sets that perform poorly at off design conditions.

3.4 Decomposition of Process Architecture

The combinatorial optimization problem in Eq. (3.45) is very expensive especially when there are large number of candidate controlled variables. If the process plant can be decomposed, then the CV selection problem can be solved independently for each section.

**Proposition:**

The combinatorics of selecting CVs under a decomposed scheme is less than the combinatorics of the original problem.
Proof:
Let the number of candidate CVs in the original problem be denoted by $N_y$ and the number of degrees of freedom be denoted by $N_u$. Let us further assume the decomposed problem has 2 partitions with candidate CVs of cardinality $N_{y1}, N_{y2}$ with $N_{y1} + N_{y2} = N_y$. Let $\binom{N_y}{j}$ denote the number of ways of choosing $j$ from $N_y$. It can easily be seen that

$$\binom{N_y}{N_u} = \sum_{j=0}^{N_u} \binom{N_{y1}}{j} \cdot \binom{N_{y2}}{N_u - j} = \left[ \sum_{j=1}^{N_u-1} \binom{N_{y1}}{j} \cdot \binom{N_{y2}}{N_u - j} \right] + \binom{N_{y1}}{N_u} + \binom{N_{y2}}{N_u} \quad (3.46)$$

Where the term in the square brackets denote the combinatorics of selection under the decomposed scheme thus it can be seen trivially that decomposition would reduce the combinatorial explosion of CV selection hence computational time. This can be easily extended to $m$ decompositions as in Eq. (3.47). Let $k_i$ unordered items be selected from $N_{yi}$ in partition $i$ with $\sum_i k_i = N_u$. Again the combinatorics under the decomposed scheme is in square brackets while the last term on the right denotes the decrease in combinatorics as compared with the overall system.

$$\binom{N_y}{N_u} = \sum_{k_1=0}^{N_u} \ldots \sum_{k_m=0}^{N_u} \prod_{i=1}^{m} \binom{N_{yi}}{k_i} = \left[ \sum_{i=1}^{m} \left( \prod_{k_i=1}^{N_u} \binom{N_{yi}}{k_i} \right) \right] + \sum_{\forall \{i \in \{1, \ldots, m\} \cap h} \sum_{\exists k_i \in \{0, \ldots, N_u\}} \left( \prod_{i=1}^{m} \binom{N_{yi}}{k_i} \right) \quad (3.47)$$

The method of process decomposition into different sections proceeds from Chapter 2. A second order nonlinear model (Eq. 2.3) is used to approximate the process data from the plant. The parameters of this model have been shown to translate to three types of connectivity amongst process subunits, namely: latent (A, H), induced (B) and extrinsic(C). Here, focus is in the connectivity relationships modulated by inputs (i.e. B). To begin estimation, prior assumptions are made on the parameters. From Bayes law, posterior densities are estimated based on likelihood densities and defined priors. Extra parameters known as hyperparameters are used to parameterize the covariance of the parameters. An expectation maximization algorithm is used to iteratively estimate the set of parameters $\{A, B, C, H\}$ of the model. From these estimates, connectivity information is drawn and classified as weak or strong depending on a user defined threshold $\zeta$ as shown in Eq. (3.48), where $B_{ik}^j$ is the bilinear parameter between variables/units $i$ and $k$ modulated by input $u_j$. If the connectivity between all variables of any two variables/units exceed the
threshold, then they are considered together as a section. For the purpose of decomposing into sections, different cases will be considered for examination, ranging from the full consideration of the process as one indivisible unit for the purpose of CV selection to the consideration of each unit operation to be disparate from one another. Each of this different cases would be considered for CV selection in terms of optimality, runtime and feasibility. Following selection criteria were applied:

\[
\text{connectivity}(k \to i) \triangleq \begin{cases} \text{strong if } B_{ik}^l / B_{ii}^l \geq \zeta \\ \text{weak otherwise} \end{cases} \tag{3.48}
\]

Given a set of islands denoted by \( k = 1, 2, ..., K \), the selection matrices for each island \( k \) is denoted by \( H_{y,k}, H_{u,k}, H_{d,k} \) for the candidate controlled variables, the manipulated variables and the disturbances respectively. The following are easily derived:

\[
J_{uu}^k = H_{u,k} J_{uu}^k H_{u,k}^T, \quad J_{ud}^k = H_{u,k} J_{ud}^k H_{d,k}^T, \quad G^k = H_{y,k} G, \quad G_d^k = H_{y,k} G_d H_{d,k}^T, \quad W_n^k = H_{y,k} W_n H_{y,k}^T, \quad W_d^k = H_{y,k} W_d H_{d,k}^T.
\]

Thus for each island \( P_k \) represents the logical vector denoting the candidate variables chosen from island \( k \).

The optimization in Eq. (3.45) becomes

\[
\text{for } k = 1, 2, ..., K \\
\min_{P_k} [L^k, J_c^k] \\
\text{Subject to} \\
L = \frac{1}{6(N_u^k + N_d^k)} \left\| (J_{uu}^k)^{1/2}(G^k)^{-1} \left[ (G^k (J_{uu}^k)^{-1} J_{ud}^k - G_d^k) W_n^k W_d^k \right] \right\|_F^2 \\
J_c^k(c) = \sigma^{-1}(G^k) \\
\sum_{i \in \Pi^k_m} F_i = \xi_m^k \forall m \in 1, ..., M^k \\
t_d(u_i, y_j) \leq \chi_j
\]

**Proposition:**

For a completely unconnected system, the CV selection problem for the original problem collapses to the CV selection problem of the decomposed system.
Proof

First, let a process be given such that it can be partitioned into two sections as follows:

\[ y = \{y_1, y_2\}, u = \{u_1, u_2\}, d = \{d_1, d_2\} \text{ such that } y_1 \in \mathbb{R}^{Ny_1}, y_2 \in \mathbb{R}^{Ny_2}, u_1 \in \mathbb{R}^{Nu_1}, u_2 \in \mathbb{R}^{Nu_2}, d_1 \in \mathbb{R}^{Nd_1}, d_2 \in \mathbb{R}^{Nd_2}. \]

Additionally, assume \( u_1, y_1, d_1 \) have no interaction with \( u_2, y_2, d_2 \).

Therefore the following can be written.

\[
G = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}, \quad G_d = \begin{bmatrix} G_{d1} & 0 \\ 0 & G_{d2} \end{bmatrix}, \quad J_{uu} = \begin{bmatrix} J_{uu,1} & 0 \\ 0 & J_{uu,2} \end{bmatrix}, \quad J_{ud} = \begin{bmatrix} J_{ud,1} & 0 \\ 0 & J_{ud,2} \end{bmatrix}
\]

Ignoring the scaling factor, the loss function for the whole system can be written as

\[
L = \left\| \frac{1}{2} \left( J_{uu}^{-1} J_{ud} - G^{-1} G_d \right) W_d J_{uu}^2 G^{-1} W_n \right\|_F^2
\]  

(3.50)

\[
= \left\| \frac{1}{2} \left( J_{uu,1}^2 G_{11}^{-1} - G_{d1} \right) W_{d1} \begin{bmatrix} J_{uu,1}^2 G_{11}^{-1} 0 \\ 0 & J_{uu,2}^2 G_{22}^{-1} \end{bmatrix} \begin{bmatrix} G_{d1}^{-1} J_{ud,1} W_{d1} & 0 \\ 0 & (G_{d2} J_{uu,2}^{-1} J_{ud,2} - G_{d2}) W_{d2} \end{bmatrix} \right\|_F^2
\]  

(3.51)

\[
= \left\| \begin{bmatrix} J_{uu,1}^2 G_{11}^{-1} & 0 \\ 0 & J_{uu,2}^2 G_{22}^{-1} \end{bmatrix} \begin{bmatrix} G_{d1}^{-1} J_{ud,1} W_{d1} & 0 \\ 0 & (G_{d2} J_{uu,2}^{-1} J_{ud,2} - G_{d2}) W_{d2} \end{bmatrix} \right\|_F^2
\]  

(3.52)

\[
= \left\| \begin{bmatrix} J_{uu,1}^2 G_{11}^{-1} & 0 \\ 0 & J_{uu,2}^2 G_{22}^{-1} \end{bmatrix} \begin{bmatrix} G_{d1}^{-1} J_{ud,1} W_{d1} & 0 \\ 0 & (G_{d2} J_{uu,2}^{-1} J_{ud,2} - G_{d2}) W_{d2} \end{bmatrix} \right\|_F^2
\]  

(3.53)

\[
= \left\| \begin{bmatrix} J_{uu,1}^2 G_{11}^{-1} & 0 \\ 0 & J_{uu,2}^2 G_{22}^{-1} \end{bmatrix} \begin{bmatrix} G_{d1}^{-1} J_{ud,1} W_{d1} & 0 \\ 0 & (G_{d2} J_{uu,2}^{-1} J_{ud,2} - G_{d2}) W_{d2} \end{bmatrix} \right\|_F^2
\]  

(3.54)

\[
= \left\| \begin{bmatrix} J_{uu,1}^2 G_{11}^{-1} & 0 \\ 0 & J_{uu,2}^2 G_{22}^{-1} \end{bmatrix} \begin{bmatrix} G_{d1}^{-1} J_{ud,1} W_{d1} & 0 \\ 0 & (G_{d2} J_{uu,2}^{-1} J_{ud,2} - G_{d2}) W_{d2} \end{bmatrix} \right\|_F^2
\]  

(3.55)

\[
L = L_1 + L_2
\]  

(3.56)

This can be extended to any number of partitions by mathematical induction of the first kind hence the proof.
3.5 Optimization Scheme

3.5.1 Multiagent Optimization

Large scale combinatorial optimization problems such as posed in Eq. (3.45, 3.49) by nature do not possess convexity and/or have discontinuous search space therefore traditional methods of optimization fail considerably (Gebreslassie and Diwekar, 2016). Additionally, the exhaustive nature of the branch and bound renders it unsuitable for large scale problems and/or online applications thus a metaheuristic approach is more suitable.

The multiagent framework used in this work employs metaheuristic optimization strategies such as the Efficient ant colony algorithm (Dorigo, 1991), Simulated annealing (Kirkpatrick, 1984) and the genetic algorithm (Hayes-Roth, 1975). As these algorithms require careful design to guarantee global optimality due to their random initialization and random search procedures, the multiagent framework allows for the combination of these stochastic algorithms and procedures into one framework. This framework supports the cooperation search by a group of algorithmic agents which are connected through the frameworks predefined information sharing protocol. By using several agents, the strengths of each agent can be exploited. Similar to the coordination of the biological organism by the central nervous system, each of the islands would be coordinated by the MAOP for the purpose of solving Eq. (3.49). The results of each island would then be aggregated as the solution for the whole system.

As shown in Fig 3.1, the Multiagent Optimization (MAOP) framework includes the following: representation of the problem to be solved (this involves definition of objective function and constraints), the global sharing memory environment, pool of algorithmic solvers (agents), scheduler that allocates resource and the execution of the algorithmic agents to solve the assigned task(s), processing and retrieval of final solutions.

![Diagram](image)

**Figure 3.1** The general framework of the multiagent platform
The solution strategy for an optimization problem in this framework involves solution sharing. This is predicated on a specific communication protocol between each agent and the global sharing environment. Every agent is unique and consists of an algorithmic procedure and a communication protocol. Agents employ the information gained from the algorithm to update the global sharing memory with better solutions until convergence and/or termination of iterations. The heterogeneous multiagent framework promises faster runtime compared to branch and bound. Fig. 3.2 shows a flowchart representation of a single agent.

![Flowchart representation of a single agent](image)

**Figure 3.2** An agent in MAOP framework (Gebreslassie and Diwekar, 2016).

The multiagent algorithm begins by initializing OPTIONS and PARAMETERS for the global algorithm. PARAMETERS include the number, type of agents and framework ID while OPTIONS includes the maximum number of iterations for the global algorithm (which calls the individual agents) $MaxIter$ and the maximum number of consecutive iterations $ConIter$ with no objective function improvement. This proceeds with initializing the global sharing memory environment (initial solutions from each algorithmic agent is used to obtain a solution archive). Then each agent is also initialized with its own specific parameters (e.g. Population size for Efficient Genetic Algorithm). Once the global algorithm proceeds, at every iteration, a call is made to each agent in a random manner and the solution from the previous agents which has been communicated to the global sharing environment and duly updated is provided as an initialization for the next agent. This is stored in the local memory of each agent being called. Each agent then makes its own call to solve the objective function posed in Eq. (3.49) for each section as partitioned. After processing all sections, the results are aggregated to form the indices of the CV selected including the value
of the loss function and the controllability. The agents updates global solution (CV indices and current value of the objective function. until a termination criteria (MaxIter or ConIter) is satisfied.

3.6 Case study: Acid gas removal (AGR) Unit

3.6.1 Process Description

This section entails the implementation of the proposed algorithm on an IGCC power plant case study based upon the model developed by (Bhattacharyya et al., 2011). This process consists of more than 20 unit operations, 5 recycles and 16 degrees of freedom. Table 3.1 enumerates all candidate controlled variables for the acid gas removal unit.

Table 3.1 List of all candidate primary controlled variables

<table>
<thead>
<tr>
<th>S/n</th>
<th>Candidate controlled variable</th>
<th>Number of variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Lean solvent (Selexol ) flow rate</td>
<td>1</td>
</tr>
<tr>
<td>2.</td>
<td>Liquid phase CO₂ fraction in CO₂ absorber</td>
<td>15</td>
</tr>
<tr>
<td>3.</td>
<td>Liquid phase H₂S fraction in CO₂ absorber</td>
<td>15</td>
</tr>
<tr>
<td>4.</td>
<td>Vapor phase CO₂ fraction in CO₂ absorber</td>
<td>15</td>
</tr>
<tr>
<td>5.</td>
<td>Vapor phase H₂S fraction in CO₂ absorber</td>
<td>15</td>
</tr>
<tr>
<td>6.</td>
<td>Temperature of stages in CO₂ absorber</td>
<td>15</td>
</tr>
<tr>
<td>7.</td>
<td>Semilean solvent flow rate</td>
<td>1</td>
</tr>
<tr>
<td>8.</td>
<td>Liquid phase CO₂ fraction in H₂S absorber</td>
<td>23</td>
</tr>
<tr>
<td>9.</td>
<td>Liquid phase H₂S fraction in H₂S absorber</td>
<td>23</td>
</tr>
<tr>
<td>10.</td>
<td>Vapor phase CO₂ fraction in H₂S absorber</td>
<td>23</td>
</tr>
<tr>
<td>11.</td>
<td>Vapor phase H₂S fraction in H₂S absorber</td>
<td>23</td>
</tr>
<tr>
<td>12.</td>
<td>Temperature of stages in H₂S absorber</td>
<td>23</td>
</tr>
<tr>
<td>13.</td>
<td>H₂ Recovery Vessel Pressure</td>
<td>1</td>
</tr>
<tr>
<td>14.</td>
<td>Medium Pressure Vessel Pressure</td>
<td>1</td>
</tr>
<tr>
<td>15.</td>
<td>H₂ recovery H₂ vapor fraction</td>
<td>1</td>
</tr>
<tr>
<td>16.</td>
<td>H₂ recovery CO₂ vapor fraction</td>
<td>1</td>
</tr>
<tr>
<td>17.</td>
<td>H₂ recovery H₂ liquid fraction</td>
<td>1</td>
</tr>
<tr>
<td>18.</td>
<td>H₂ recovery CO₂ liquid fraction</td>
<td>1</td>
</tr>
<tr>
<td>19.</td>
<td>H₂ recovery temperature</td>
<td>1</td>
</tr>
<tr>
<td>20.</td>
<td>Medium pressure H₂ vapor fraction</td>
<td>1</td>
</tr>
<tr>
<td>21.</td>
<td>Medium pressure CO₂ vapor fraction</td>
<td>1</td>
</tr>
<tr>
<td>22.</td>
<td>Medium pressure H₂ liquid fraction</td>
<td>1</td>
</tr>
<tr>
<td>23.</td>
<td>Medium pressure CO₂ liquid fraction</td>
<td>1</td>
</tr>
<tr>
<td>24.</td>
<td>Stripping syngas flow</td>
<td>1</td>
</tr>
<tr>
<td>25.</td>
<td>Liquid phase CO₂ fraction in H₂S concentrator</td>
<td>5</td>
</tr>
<tr>
<td>26.</td>
<td>Liquid phase H₂S fraction in H₂S concentrator</td>
<td>5</td>
</tr>
<tr>
<td>27.</td>
<td>Vapor phase CO₂ fraction in H₂S concentrator</td>
<td>5</td>
</tr>
<tr>
<td>28.</td>
<td>Vapor phase H₂S fraction in H₂S concentrator</td>
<td>5</td>
</tr>
<tr>
<td>29.</td>
<td>Temperature of stages in H₂S concentrator</td>
<td>5</td>
</tr>
</tbody>
</table>
### 3.6.2 Decomposition of AGR Unit

The acid gas removal unit is simulated in DYNSIM® (http://software.schneider-electric.com/). The model is simulated from steady state with a 20% deviation in CO₂-laden syngas flowrate (278,732.281kgh⁻¹ – 337,478.732kgh⁻¹) and (278,732.281kgh⁻¹ – 222,985.825kgh⁻¹). 1000 data points in the interval of 0.25s are collected between time \( t = 0 \) and \( t = 250 \) s. These output variables are collected in the DYNSIM interface and exported to MATLAB where they are normalized and preprocessed (section 2.4.3). The obtained data from the high fidelity nonlinear model is then approximated by a second order nonlinear model in Eq. (2.3) employed in the system identification scheme. In this scheme, the parameters are then estimated through a dual expectation maximization scheme with Bayesian inferencing. Bayesian inferencing suffices here due to the use of informative priors. These priors allow us impose values that would prevent unrealistic connectivity results e.g. unconnected units should have a connectivity of zero. Upon convergence of the identification scheme, the parameters are post processed according to Eq. (3.48) to identify the variables which are strongly connected and hence grouped together into sections.

Given the notation as shown for various units:

1. CO₂ absorber \( (T_1) \)
2. H₂S absorber \( (T_2) \)
3. Selexol stripper \( (T_3) \)
4. H₂S concentrator \( (T_4) \)
5. H₂ recovery K.O drum \( (D_1) \)
6. H₂ recovery drum \( (D_2) \)
7. Medium pressure flash drum \( (D_3) \)
8. Low pressure flash drum \( (D_4) \)

The hardware is decoupled into five different configurations (depending on the connectivity threshold \( \zeta \)) as shown in Table 3.2.

#### Table 3.2 Decomposition of the AGR unit into sections.

<table>
<thead>
<tr>
<th>s/n</th>
<th>Connectivity threshold</th>
<th>Decomposition groups</th>
<th>Number of sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>( \zeta = \infty )</td>
<td>([T_1], [T_2], [T_3], [T_4], [D_1], [D_2], [D_3], [D_4])</td>
<td>8</td>
</tr>
<tr>
<td>2.</td>
<td>( \zeta = 100 )</td>
<td>([T_1], [T_2], [T_3], [T_4], [D_1], [D_2], [D_3], [T_4])</td>
<td>5</td>
</tr>
<tr>
<td>3.</td>
<td>( \zeta = 10 )</td>
<td>([T_1], [T_2], [T_3], [T_4], [D_1], [D_2], [D_3])</td>
<td>3</td>
</tr>
<tr>
<td>4.</td>
<td>( \zeta = 1 )</td>
<td>([T_1], [T_2], [T_3], [T_4], [D_1], [D_2], [D_4])</td>
<td>2</td>
</tr>
<tr>
<td>5.</td>
<td>( \zeta = 0.1 )</td>
<td>([T_1], [T_2], [T_3], [T_4], [D_1], [D_2], [D_3], [D_4])</td>
<td>1</td>
</tr>
</tbody>
</table>
3.6.3 Identification of Candidate Sets of Controlled Variables

3.6.3.1 Objective Function

To implement the controlled variable selection, first, an operational objective function is defined. Similar to (Bankole et al., 2018a), this objective function takes into account the operational cost of the acid gas removal unit which includes cost of utilities, feeds, waste streams, products and energy generation. The utilities encompasses the following: ammonia refrigeration duties, compressor power, heating and cooling costs. Feeds to the unit include shifted syngas, exit tail gas sent to the Claus unit, and makeup solvent used to replenish lost solvent in the system. Waste streams considered includes carbon monoxide and hydrogen unrecovered by the separation unit thus constitute losses and/or inefficiency. The contribution of this is quantified by how much power can be obtained in the turbine if these constituted part of the recovered syngas. Negative cost was attributed to energy recovered from the gas turbine from the clean syngas. All costs were determined and normalized by converting to an equivalent electrical cost. This is done by determining the electrical power required for pumping and the compressor power required to generate heat duty. The assumed cost of electricity is $0.0943/kWh. Make up solvent cost is obtained from (Bucklin and Schendel, 1984). Lastly, product losses such as hydrogen and carbon monoxide which are valuable are accounted for. The entire cost function is given below:

$$ J \left( \frac{\$/h}{W} \right) = 6.28 \left( m_{makeup solvent} \right) + 0.00982 \left( m_{steam} \right) + 17.9 \sum_{k} \dot{Q}_k + 13.7 \sum_{l} \dot{Q}_l $$

$$ + 0.836 \sum_{m} \dot{Q}_m + 94.3 \left( \sum_{i} \dot{W}_i + \sum_{j} \dot{W}_j \right) + 3.03 \left( \dot{m}_{H_2} + \dot{m}_{CO} \right) \left( 3.57 \right) $$

(3.57)

Units of power ($\dot{W}$), heat duty ($\dot{Q}$) and mass flow rate ($m$) are given as MW, MW and kg/h.

Given that the set of active constraints are controlled with the pairings generated for the process as outlined in (Jones et al., 2014) it is sought to select additional degrees of freedom for control as the AGR process has five degrees of freedom left for use. Therefore the remaining degree of freedom which span the unconstrained space (as described in section 3.3.2) are to be implemented in controlling five additional controlled variables from 230 leading to \( \binom{230}{5} = 5.14 \times 10^9 \) alternative combinations. The corresponding hardware for the manipulated variables, both spanning the constrained and unconstrained space are given in Table 3.3.
Table 3.3 List of manipulated variables and hardware for the AGR unit.

<table>
<thead>
<tr>
<th>S/n</th>
<th>Manipulated variable</th>
<th>Hardware</th>
<th>Active constraint/Inactive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>LP flash vessel pressure</td>
<td>LP CO(_2) compressor</td>
<td>Active</td>
</tr>
<tr>
<td>2.</td>
<td>MP flash vessel pressure</td>
<td>MP CO(_2) compressor</td>
<td>Inactive</td>
</tr>
<tr>
<td>3.</td>
<td>H(_2) recovery vessel pressure</td>
<td>H(_2) recovery compressor</td>
<td>Inactive</td>
</tr>
<tr>
<td>4.</td>
<td>Speed control</td>
<td>H(_2) compressor pressure</td>
<td>Active</td>
</tr>
<tr>
<td>5.</td>
<td>Recovered acid gas temperature</td>
<td>Cooling water flowrate to heat exchanger</td>
<td>Active</td>
</tr>
<tr>
<td>6.</td>
<td>Stripper reboiler duty</td>
<td>Steam flow rate to reboiler</td>
<td>Active</td>
</tr>
<tr>
<td>7.</td>
<td>Stripping syngas flow rate</td>
<td>Stripping syngas flow rate to H(_2)S concentrator</td>
<td>Inactive</td>
</tr>
<tr>
<td>8.</td>
<td>H(_2)S concentrator pressure</td>
<td>Stripped gas compressor</td>
<td>Active</td>
</tr>
<tr>
<td>9.</td>
<td>Semi lean solvent flow rate</td>
<td>Semilean flowrate control valve</td>
<td>Inactive</td>
</tr>
<tr>
<td>10.</td>
<td>H(_2) cooler duty</td>
<td>Cooling water flowrate to heat exchanger</td>
<td>Active</td>
</tr>
<tr>
<td>11.</td>
<td>Lean solvent flow rate</td>
<td>Control valve regulating solvent flow to CO(_2) absorber</td>
<td>Inactive</td>
</tr>
</tbody>
</table>

The degree of freedom analysis is as follows: the H\(_2\) recovery flash pressure, the semi-lean solvent flow and the lean solvent flow are available for the CO\(_2\) absorber, while only the lean solvent flow is available for the H\(_2\)S absorber, similarly, the Stripping syngas flow rate is available for control of the H\(_2\)S concentrator. Finally the Medium pressure and the low pressure CO\(_2\) flash vessels have their respective flash pressure as degrees of freedom. This analysis is necessary for the subset selection constraint \(\sum_{i \in \Pi_i} F_i = \xi_m \forall m \in 1, ..., M^K\) formulated in Eq. (3.49). The disturbances considered include upstream variation in syngas flow and composition from the gasification section. For CV selection of the AGR process, only five unit operations have degrees of freedom for control namely \(T_1, T_2, T_4, D_2, D_3\). Hence the decomposition in Table 3.4.

Table 3.4 Decomposition of the AGR unit into sections for CV selection.

<table>
<thead>
<tr>
<th>s/n</th>
<th>Connectivity threshold</th>
<th>Decomposition groups</th>
<th>Number of sections</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>(\zeta = \infty)</td>
<td>([T_1], [T_2], [T_4], [D_2], [D_3])</td>
<td>5</td>
</tr>
<tr>
<td>2.</td>
<td>(\zeta = 100)</td>
<td>([T_1, T_2], [D_2], [D_3], [T_4])</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>(\zeta = 10)</td>
<td>([T_1, T_2, T_4], [D_2], [D_3])</td>
<td>3</td>
</tr>
<tr>
<td>4.</td>
<td>(\zeta = 1)</td>
<td>([T_1, T_2, T_4, D_2], [D_3])</td>
<td>2</td>
</tr>
<tr>
<td>5.</td>
<td>(\zeta = 0.1)</td>
<td>([T_1, T_2, T_4, D_2, D_3])</td>
<td>1</td>
</tr>
</tbody>
</table>
3.6.3.2 Selection of Pareto sets with Multiagent Optimization

The optimization as defined in Eq. (3.49) was implemented on an Intel® Xeon® CPU E-5-1620 v2 with 32GB RAM using the heterogeneous multiagent framework with three agents: efficient ant colony optimization, simulated annealing and genetic algorithm. Secondly for comparison, the optimization is solved using a stand-alone branch and bound and parallelized branch and bound as described in (Jones et al., 2014). The parallelized branch and bound is deployed on a MATLAB® distributed computing platform with 54 workers. The heterogeneous multiagent optimization framework is programmed in MATLAB. A framework ID is allocated to the MAOP solver which indicates which agents are to be utilized in the solution. Each agent is initialized with local parameter settings and are only accessed by the agents. Contrarily, global parameters are accessed by all the agents in the memory sharing environment. The termination criteria for the framework is the maximum global iteration $Maxiter$ and/or the global tolerance $Eps$ which is the minimum allowable difference between any two consecutive solutions within a fixed number of consecutive iterations denoted as $ConIter$. The termination criteria for the local agents follow a similar approach. The parameters for the agents are given in Table 3.4. For simulated annealing, the objective cost (referred to as system energy) is minimized by accepting solutions from random perturbations of previous solution states (mimicking particle motion in annealing). If the new objective function is lower, then the new state is accepted otherwise it is accepted according to a probability function similar to the Boltzmann distribution function. Initially, probability of accepting new solutions (whether worse or better) is initially high (i.e. at the initial temperature) and consequently drops so that only solutions that minimize the objective function are accepted. This is controlled by the ‘quenching factor’. This continues until the iteration proceeds to reach maximum iteration (stop temperature). Extensive details of this algorithm is provided in (Kim and Diwekar, 2002).

For the ant colony optimization, each random move from one a state to the next is denoted by an ‘edge’ and at each iteration, the solution archive which consists of the best solutions is populated. Each edge has a pheromone level attached to it which depends on the quality of solution. Again, similar to the simulated annealing, the pheromone evaporation factor controls the quality of solutions that are accepted as iterations proceed. Extensive details can be found in (Gebreslassie and Diwekar, 2015)
Finally, the efficient genetic algorithm works by mutation of solutions. At each iteration, a fraction of the best solutions are selected for mutation to improve the objective function. Better solutions are retained and used for crossover and mutation while poorer solutions are discarded. All the agents are cast into the multiagent framework with parameters as described in Table 3.5. The results of the execution time compared with parallelized branch and bound\(^2\) and conventional branch and bound are as given in Fig 3.3. Only the cases which correspond to 1,2 and 3 islands are shown are the rest are trivially fast in comparison.

**Table 3.5 Parameters for the agents**

<table>
<thead>
<tr>
<th>Simulated Annealing Agent</th>
<th>Initial temperature</th>
<th>Quenching factor</th>
<th>Maximum success</th>
<th>MaxIter</th>
<th>Maximum consecutive rejection</th>
<th>Stop temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>0.9</td>
<td>20</td>
<td>200</td>
<td>30</td>
<td>1e-6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genetic Algorithm</th>
<th>Population</th>
<th>Mutation rate</th>
<th>Selection</th>
<th>MaxIter</th>
<th>ConIter</th>
<th>Eps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>0.0075</td>
<td>0.55</td>
<td>1000</td>
<td>20</td>
<td>1e-5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Efficient ant colony</th>
<th>Number of ants</th>
<th>Pheromone evaporation</th>
<th>Solution Archive</th>
<th>MaxIter</th>
<th>ConIter</th>
<th>Eps</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>0.7</td>
<td>50</td>
<td>2000</td>
<td>10</td>
<td>1e-5</td>
</tr>
</tbody>
</table>

\(^2\) Comparison is performed with Jones et al 2014 as the algorithm developed there is formulated with selection constraints and handles controllability.
Figure 3.3 Running time for BB (branch and bound), PBB (Parallelized branch and bound) and MAOP (Multiagent optimization) for the cases of 1, 2 and 3 islands respectively.

The results of the MAOP differs from island to island and as expected, with more decomposition (i.e. more islands) comes an increased in execution speed but reduced accuracy of the best set of controlled variables. For the case with only one island i.e. considering the process as a whole, The MAOP offers approximately 90% reduction in execution time in comparison to the standalone branch and bounds and 70% reduction in comparison to parallelized branch and bound. This gains could be higher if the MAOP is parallelized as well. In this work, the agents are run sequentially. For brevity, Table 3.5 shows the best three global optimum set of controlled variable set for each decomposition (from all agents). In the first row, 5 distinct islands are considered which corresponds to a connectivity threshold of $\zeta = \infty$ as in Table 3.2. Similar correspondence between the connectivity threshold for other rows in Table 3.5 and Table 3.2 holds. The first entry with 5 islands indicates controlled variable selection was carried out for each unit operation separately and merged while the last entry considers the whole process as a unit aggregate and corresponds to traditional controlled variable selection.
<table>
<thead>
<tr>
<th>Islands</th>
<th>Decomposition groups</th>
<th>Global optimum set of CVs</th>
<th>Average Loss ($h^{-1}$)</th>
<th>Controllability $\sigma^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.</td>
<td>• [CO₂ absorber]</td>
<td>3 11 16 22 33</td>
<td>78.215</td>
<td>62.63</td>
</tr>
<tr>
<td></td>
<td>• [H₂ recovery drum]</td>
<td>6 9 17 26 32</td>
<td>79263.6</td>
<td>10.83</td>
</tr>
<tr>
<td></td>
<td>• [MP flash drum]</td>
<td>4 10 20 23 30</td>
<td>179073.5</td>
<td>10.72</td>
</tr>
<tr>
<td>4.</td>
<td>• [CO₂ absorber; H₂S absorber]</td>
<td>1 13 21 24 27</td>
<td>30.09</td>
<td>9.93</td>
</tr>
<tr>
<td></td>
<td>• [H₂ recovery drum]</td>
<td>5 7 16 22 33</td>
<td>350.8</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>• [MP flash drum]</td>
<td>5 13 18 25 29</td>
<td>5520</td>
<td>5.41</td>
</tr>
<tr>
<td>3.</td>
<td>• [CO₂ absorber, H₂S absorber, H₂S concentrator]</td>
<td>5 8 16 22 31</td>
<td>30.195</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>• [H₂ recovery drum]</td>
<td>5 14 18 25 31</td>
<td>205.392</td>
<td>1.03</td>
</tr>
<tr>
<td></td>
<td>• [MP flash drum]</td>
<td>1 14 21 24 31</td>
<td>24853.3</td>
<td>4.17</td>
</tr>
<tr>
<td>2.</td>
<td>• [CO₂ absorber, H₂S absorber, H₂ recovery drum, H₂S concentrator]</td>
<td>5 15 21 24 28</td>
<td>3.78</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td>• [MP flash drum]</td>
<td>1 15 21 25 28</td>
<td>30.179</td>
<td>1.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 12 21 22 34</td>
<td>291.42</td>
<td>0.32</td>
</tr>
<tr>
<td>1.</td>
<td>• [CO₂ absorber, H₂ recovery drum, H₂S absorber, H₂S concentrator, MP flash drum]</td>
<td>2 12 19 26 34</td>
<td>1.488</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 11 19 26 34</td>
<td>1.723</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 11 19 26 34</td>
<td>1.745</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Figs. 3.4-3.5 shows the best economic loss and controllability obtained as a function of the number of islands in the decomposition for both the average case and the worst case scenario.

---

3 CV corresponding to Indices are detailed in Table 3.7
Firstly it can be observed that both the average and worst case economic loss potentially increase as the number of sections employed increases. This is due to the loss in information between the different sections which occurs as a result of decomposition and separation. Alternatively, the decomposition can be viewed as enforcing an assumption of non-interaction between the considered sections. Therefore this translates to a concomitant increase in the economic loss derived due to the decomposition. It can be observed from Fig. 3.4 that the case where the process is decomposed into two (2) sections, the optimal value of the loss function is close to the loss function of the case without any decomposition (i.e. the case denoted by no. of sections as 1).
Furthermore, it is observed that the CV sets that perform best for the average case loss equally perform best for the worst case loss thus the controllability plots for both cases are one and same and is plotted in Fig. 3.5. With respect to the sections, the controllability measure similarly decreases implying ease of control with less sections. It can be seen that the controllability measure increases by several orders of magnitude from the case of one section ($\sigma^{-1} = 0.0005$) where the process is considered as a whole to the extreme case where all unit operations are considered as separate sections ($\sigma^{-1} = 1$). It is observed that the loss of controllability is minimal until when the no of sections is between 2 or 3. Considering the tradeoffs, partitioning of the acid gas removal unit into two sections appears very attractive since the loss in the economic objective (worst case: $43.62\text{h}^{-1}$, average case: $3.78\text{h}^{-1}$) and controllability is minimal, while resulting in about 40% improvement in computational time on average for all three methods (BB, PBB, MAOP). The results of using two sections are given in Table 3.6 and corresponds to the following sets of CVs [Stage 15 liquid phase H$_2$S concentration in CO$_2$ absorber, Stage 2 Temperature in H$_2$S absorber, H$_2$ recovery flash pressure vapor H$_2$ fraction, Medium pressure flash liquid H$_2$ fraction, Stage 4 liquid phase H$_2$S concentration in H$_2$S concentrator]
### Table 3.7 Indices of candidate controlled variables

<table>
<thead>
<tr>
<th>Index</th>
<th>Controlled variable*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>( (x\ CO_2)_{04} )</td>
</tr>
<tr>
<td>2.</td>
<td>( (xH_2S)_{07} )</td>
</tr>
<tr>
<td>3.</td>
<td>( (xH_2S)_{13} )</td>
</tr>
<tr>
<td>4.</td>
<td>( (xH_2S)_{14} )</td>
</tr>
<tr>
<td>5.</td>
<td>( (xH_2S)_{14} )</td>
</tr>
<tr>
<td>6.</td>
<td>( T_{14} )</td>
</tr>
<tr>
<td>7.</td>
<td>( (xCO_2)_{04} )</td>
</tr>
<tr>
<td>8.</td>
<td>( (xCO_2)_{23} )</td>
</tr>
<tr>
<td>9.</td>
<td>( (xH_2S)_{09} )</td>
</tr>
<tr>
<td>10.</td>
<td>( (xH_2S)_{19} )</td>
</tr>
<tr>
<td>11.</td>
<td>( (yCO_2)_{01} )</td>
</tr>
<tr>
<td>12.</td>
<td>( (yCO_2)_{19} )</td>
</tr>
<tr>
<td>13.</td>
<td>( (yH_2S)_{20} )</td>
</tr>
<tr>
<td>14.</td>
<td>( T_{01} )</td>
</tr>
<tr>
<td>15.</td>
<td>( T_{02} )</td>
</tr>
<tr>
<td>16.</td>
<td></td>
</tr>
<tr>
<td>17.</td>
<td></td>
</tr>
<tr>
<td>18.</td>
<td></td>
</tr>
<tr>
<td>19.</td>
<td></td>
</tr>
<tr>
<td>20.</td>
<td></td>
</tr>
<tr>
<td>21.</td>
<td></td>
</tr>
<tr>
<td>22.</td>
<td></td>
</tr>
<tr>
<td>23.</td>
<td></td>
</tr>
<tr>
<td>24.</td>
<td></td>
</tr>
<tr>
<td>25.</td>
<td></td>
</tr>
<tr>
<td>26.</td>
<td></td>
</tr>
<tr>
<td>27.</td>
<td>( (xCO_2)_{02} )</td>
</tr>
<tr>
<td>28.</td>
<td>( (xH_2S)_{04} )</td>
</tr>
<tr>
<td>29.</td>
<td>( (xH_2S)_{05} )</td>
</tr>
<tr>
<td>30.</td>
<td>( (yCO_2)_{05} )</td>
</tr>
<tr>
<td>31.</td>
<td>( (yH_2S)_{03} )</td>
</tr>
<tr>
<td>32.</td>
<td>( (yH_2S)_{05} )</td>
</tr>
<tr>
<td>33.</td>
<td>( T_{02} )</td>
</tr>
<tr>
<td>34.</td>
<td>( T_{03} )</td>
</tr>
</tbody>
</table>

*denotes stage number
3.6.3.3 Posteriori Analysis

Upon the generation of the Pareto optimal set of candidate CVs for each decomposition case, it remains expedient to examine the performance of the controlled variables at off design conditions. This was carried out (Jones et al., 2014) by using an equation solver within Aspen Plus which allows for the evaluation of the loss at fixed values of the disturbances. Here, a similar mechanism is utilized however the process data is collected from a dynamic simulator (DYNSIM) (Zitney et al., 2012) at other values of the DVs other than their nominal values at which the linearization was performed. This study shows the sensitivity of the average and/or worst case loss function with respect to sectioning the process. The disturbances considered include variation of syngas flow rate from the gasification section to the acid gas removal unit at 80, 90, 110, and 120% of nominal values at steady state. This disturbances have been simulated under the assumption that the active constraints do not change.

Fig. 3.6a shows the variation in the average economic loss at different off-design points (case for five sections not shown as average losses are about one order of magnitude higher for all values of disturbances examined), the increase in loss function as the number of sections increase is evident, which is aligned with the results of Fig 3.4. It is clear that the nominal case presents the lowest loss irrespective of the number of decompositions. The results clearly shows the high nonlinearity of the system being studied. It is observed that there is significant increase in the economic loss with increased no of sections when the plant throughput is increased to 110%. Interestingly, the case for 120% increase has a higher economic loss than the 110% case for the undecomposed plant case, but as the number of sections increase, the deviation in average loss becomes lower compared to the 110% case examined. Similar nonlinearity can be observed when the 80% throughput case is compared with the case with 90% throughput.

Similar results can be seen in examining how the controllability changes at off design conditions as a function of the number of sections. However, it can be observed that the case with two sections still results in a reasonable compromise even while considering these off-design conditions.
Figure 3.6 Average Loss (left) and controllability (right) as a function of variation in input syngas flow rate to the H₂S absorber

3.7 Conclusion

The main goal of this chapter was to design algorithmic methods to improve the execution speed of the controlled variable selection algorithm in literature. The methodology proposed in this chapter evidenced the performance improvement of multiagent optimization technique over traditional branch and bound on the speed of execution of the controlled variable selection for processes with numerous candidate controlled variable sets. In addition, rather than consider the process in a holistic manner for the purpose of controlled variable selection, connectivity strength amongst different variables in the process is employed to decompose a process into different islands based on a user defined threshold of connectivity strength. Strongly coupled variables and unit operations are considered to be in the same islands and vice versa. The controlled variable selection algorithm is then deployed on each island in parallel and the results from each island are merged together. This decomposition however incurs sub-optimality of the loss obtained. However, our results for the AGR unit shows that two island decomposition gives a fair compromise between speed and accuracy. Furthermore, the impact of the decomposition was examined on various cases when the process shifts from nominal conditions and the losses are found to increase as the process deviates from nominal conditions, however no clear correlation was found from the decomposition and islanding on the increment of the loss.
Chapter 4

4 Optimal Control Structure Design for Cyber-Physical Systems

4.1 Background

Hybrid gas turbine-fuel cell systems exhibit immense potential for unparalleled electrical power generation efficiency with clean emissions compared to fossil fueled power generation. As hybrid power systems form the prospect of advanced power generation now and in the near future, it becomes necessary to provide a methodical control structure design for the operation of such systems. Traditionally, process experience/heuristics have sufficed for this task. For a cyber physical system (CPS) however with virtual components retrofitted/rearranged, a systematic method becomes requisite. Using a comprehensive analytical first principles based model developed from data collected at the HyPer (Hybrid Performance) facility provided by US DOE at NETL facility, a complete control structure design is embarked upon in this chapter. This approach employs a multiobjective optimization function including economics and controllability (ease of control) of the process to determine the best possible controlled variables for feedback control under varying disturbances. The discussion entails a priori analysis and heuristic based methods for prescreening, the optimization framework for selection and finally a posteriori analysis of selected variables at off design conditions. The contributions and novelties of this chapter are published in (Bankole et al., 2018b, Bankole et al., 2018c).

4.2 Introduction

As an example of a cyber-physical system, a GT-SOFC hybrid system is evaluated here for optimal CV selection. This particular system offers immense potential for superior electrical power generation efficiency (Tucker et al., 2005). Fuel cell hybrid systems can be considered as part of the polygeneration systems where integration of multiple processes are considered for coproducing multiple products such as heat, power and chemicals. These systems enable system flexibility and efficient resource utilization. By feeding the fuel cell hybrid systems with the syngas from the coal-fed integrated gasification combined cycle power systems, advanced coal based power generation with higher efficiency and cleaner emission can be accomplished (Winkler et al., 2006).
Fuel cell hybrid systems are not only attractive for stationary applications, but also for mobile systems such as ships and aircrafts.

Solid oxide fuel cells (SOFCs) (Adams et al., 2012) are small dimensional stationary, high-temperature, low-noise power generation devices with immense potential to replace currently used combustion-based power generation systems. These fuel cells mainly consist of an anode, a cathode, and a solid oxide electrolyte sandwiched inbetween. While fuel is fed to the anode, air is fed to the cathode side. These electrodes are connected externally by an electrical circuit. Extensive details can be found in literature (Bhattacharyya and Rengaswamy, 2009, Singhal and Kendall, 2003).

A GT-SOFC hybrid system has been built at the National Energy Technology Laboratory (NETL), Morgantown with the purpose of open research. This facility is part of the HYbrid PERformance (HyPer) project. The HyPer facility at NETL is a hardware simulation of a fuel cell gas turbine hybrid power system which has the capacity of reproducing power dynamics of systems in the range of 300 kW to 900 kW. The HyPer facility is a cyber physical fuel cell facility where a fuel cell model interacts with the gas turbine recuperated cycle. Other than the single-shaft gas turbine, and a high performance exhaust gas recuperator, several pressure vessels are used to capture the transient effects of the physical volumes and flow resistances of the cyber physical fuel cell, combustors, and related channelling and piping. The gas turbine is an auxiliary power unit which is a Garret Series 85 type and consists of a two stage radial compressor. The HyPer facility utilizes two recuperators with countercurrent flow to preheat air entering into the pressure vessel that facilitates to simulate the fuel cell cathode volume. The cyber system includes a real time fuel cell model that is used to control a natural gas burner which simulates the thermal output of a solid oxide fuel cell. The real time fuel cell runs on a dSpace platform, which is generally used for hardware in the loop applications.

As this technology is immensely promising, operational control of the fuel cell/gas turbine hybrid system is crucial to its development and commercialization. Thus as control is highly fundamental to the optimal and efficient operation of the system, a first step in the control structure design of the plant is to determine what variables are best for control purposes in the plant. Despite several studies on the dynamics of the HyPer facility and its interaction with the hardware (Smith et al., 2006, Winkler et al., 2006, Tucker et al., 2005), there has been no literature on the controlled
variable structure design for the plant. This necessitates the study of a complete and thorough control structure design for this cyber-physical system. A systematic approach that is relatively fast yet yields optimal CVs is pertinent for highly complex and integrated systems such as this.

4.3 Process Description

The HyPer facility is as shown in Fig. 4.1 and consists of the following subsystems:

1. Compressor/ Turbine model
2. Heat exchangers
3. Bypass valves
4. Pressure vessels (Air plenum, Combustor, Post-combustor)

A description of each subsystem is given in section 4.3.1 through 4.3.4. For brevity, all model equations for the subsystems are omitted here and can be found in (Tsai et al., 2010).

Figure 4.1 Configuration of the HyPer facility
4.3.1 Compressor/Turbine

The auxiliary power unit consists of a turbine and a compressor in a single shaft assembly capable of producing 400Hz of synchronous power. The compressor is a double stage centrifugal type compressor driven by the 120kW turbine which is encased within the compressor scroll. Exit air from the compressor exits the enclosure where concentric cooling flow is provided to the turbine inlet. The turbine nominally operates at 40,500 rpm. At this speed, approximately 2kg/s of compressed air exits the compressor at a pressure ratio of four (4)(Tucker et al., 2009).

![Figure 4.2 Compressor turbine subsystem](image)

4.3.2 Heat Recuperation

The HyPer facility consists of a combustor and air plenums which is used to reproduce the heat effluent and stack volume of the virtual 300kW SOFC. The thermal efficiency of the facility is improved by using heat exchangers (HX) to recover waste heat from the turbine exhaust to increase the temperature of the compressed air to the fuel cell stack. This closes the loop on the recuperated cycle. For the purpose of heat recovery, two parallel counter flow heat exchanger are employed. These primary heat recuperators obtain waste heat from the turbine exhaust to the compressed air which is heated up before the inlet of the SOFC cathode. This significantly increases the temperature of the compressed air thus reducing fuel requirements in the combustor. The typical effectiveness of the heat exchangers is 89% with cold side and hot side pressure losses of 2.5% and 3% respectively. The maximum temperature for both sides are given as 1150F (621°C) and 1000F (537°C) for estimated flows of 4.03lb/s (1.83kg/s) and 3.9lb/s (1.77kg/s).
4.3.3 Bypass Valves

The hardware configuration set up uses bypass valves within flow loops parallel to the mainstream flow pathways for the control of airflow to the air plenum. To minimize pressure losses in the system, no valves are used between the main pressure loop and the gas turbine. Currently, three parallel air flow control loops are being implemented in the HyPer facility, these are the Cold air (CA) bypass valve, Bleed air (BA) bypass valve and finally the Hot air (HA) bypass valve. These valves possess unique characteristics and attributes in controlling the system performance and efficiency. The bypass valves are used to mitigate the thermal management of the system, and optimize the Fuel cell-Gas turbine performance during transient operations. The bleed air valve has also been shown to increase compressor discharge pressure and to increase stall margins. The hot air valve on the other hand is effectively used to decrease cathode inlet flow. Additionally, it can lower pressure drop by 10%. Lastly the cold air valve was shown to be most influential in altering the cathode airflow, decreasing the turbine inlet temperature and increasing compressor surge margin (Tucker et al., 2005, Tucker et al., 2006).

4.3.4 Pressure Vessels

The air plenum primarily serves as a SOFC volume and piping manifold. This pressure vessel is 2.0m$^3$ in capacity. Similarly the post combustor and associated piping is a pressure vessel with a volume of 0.78m$^3$. These vessels are meant to simulate the residence time of the fuel cell. Either by use of metallic floats or apertures. The vessel and channeling is created from 2.54cm Incaloy 800AT, and is intended to work at temperatures as high as 1200K (1700ºF) at a pressure of 310kPag.
4.4 Controlled Variable Selection

This section consists of the setup of the CV selection for the HyPer facility including the a priori analysis, estimation of the variance from data obtained from the facility, and finally the cost function. Fig. 4.5 shows a flowsheet of how the model is laid out in SIMULINK, each block represents a set of equations describing the subsystems including adequate piping equations for pressure drop calculations during flue gas flow.
Firstly, a list of candidate controlled variables and a list of available manipulated variables are enumerated as shown in Tables 4.1 and 4.2. The a priori analysis consists of prescreening candidate controlled variables based on process insight and Eqs. (3.3) - (3.6). In the turbine/compressor subsection as shown in Fig. 4.5, the turbine speed is a candidate controlled variable and all other variable (including pressure and temperature of compressor and turbine exhaust) within this subsection is dependent on the turbine speed (due to the coupling of the turbine and compressor on the single shaft assembly). The electric load is a disturbance and depends on the power demand of the grid (load bank for this specific example, see Fig. 4.1). In the heat exchanger subsection, the available candidate controlled variable is the temperature to the plenum. In the air plenum, the temperature is a candidate controlled variable. Similarly, in the combustor subsystem, the temperature is a candidate controlled variable. The mass flow rate to the post-combustor depends on the hot air bypass, the cold air by pass and the mass flow rate to the plenum. The mass flow rate to the post-combustor and the mass flow rate to the plenum are both included as candidates, similarly the temperature in the post-combustor is considered as a candidate. The initial sets of candidate controlled variables reduce from 41 (Table 4.1) to 12 (Table 4.3) upon a priori analysis. This includes removal of controlled variables with poor controllability and high dead time according to Eqs. (3.5-3.6). The available degrees of freedom and disturbances are listed in Table 4.2.
Table 4.1 List of all candidate controlled variables and their respective subsystems

<table>
<thead>
<tr>
<th>s/n</th>
<th>Subsystem</th>
<th>Candidate controlled variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td>Air Mass flow rate to compressor</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td>Flue gas mass flow rate to turbine</td>
</tr>
<tr>
<td>3.</td>
<td><strong>Turbine/compressor</strong></td>
<td>Compressor Pressure</td>
</tr>
<tr>
<td>4.</td>
<td><strong>subsystem</strong></td>
<td>Compressor Temperature</td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td>Turbine Pressure</td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td>Turbine Temperature</td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td>Turbine speed</td>
</tr>
<tr>
<td>8.</td>
<td><strong>HX</strong></td>
<td>Temperature to plenum</td>
</tr>
<tr>
<td>9.</td>
<td><strong>subsystem</strong></td>
<td>Exhaust turbine temperature</td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td>Mass flow rate to heat exchanger</td>
</tr>
<tr>
<td>11.</td>
<td></td>
<td>Mass flow rate to combustor</td>
</tr>
<tr>
<td>12.</td>
<td></td>
<td>Air plenum Temperature</td>
</tr>
<tr>
<td>13.</td>
<td></td>
<td>Air plenum density</td>
</tr>
<tr>
<td>14.</td>
<td></td>
<td>Air plenum pressure</td>
</tr>
<tr>
<td>15.</td>
<td><strong>Pressure vessels</strong></td>
<td>Mass flow rate to the combustor</td>
</tr>
<tr>
<td>16.</td>
<td><strong>(Air plenum, Combustor</strong></td>
<td>Combustor temperature</td>
</tr>
<tr>
<td>17.</td>
<td><strong>Post combustor)</strong></td>
<td>Mass flow rate to Post-combustor</td>
</tr>
<tr>
<td>18.</td>
<td></td>
<td>Post-combustor Temperature</td>
</tr>
<tr>
<td>19.</td>
<td></td>
<td>Post-combustor Pressure</td>
</tr>
<tr>
<td>20.</td>
<td></td>
<td>Mass flow rate to turbine</td>
</tr>
<tr>
<td>21.</td>
<td></td>
<td>Temperature to turbine</td>
</tr>
</tbody>
</table>
Table 4.2 List of manipulated variables

<table>
<thead>
<tr>
<th>s/n</th>
<th>Subsystem</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td>Temperature from turbine</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td>Mass flow rate cold air</td>
</tr>
<tr>
<td>3.</td>
<td>Bypass valves</td>
<td>Mass flow rate hot air</td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td>Mass flow rate bleed air</td>
</tr>
</tbody>
</table>

Table 4.3 Candidate controlled variables in the hyper facility.

<table>
<thead>
<tr>
<th>Controlled variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \dot{m}_{in,pl} )</td>
<td>Mass flow rate to the plenum</td>
</tr>
<tr>
<td>( T_{in,pl} )</td>
<td>Temperature to the plenum</td>
</tr>
<tr>
<td>( T_{pl} )</td>
<td>Temperature in the plenum</td>
</tr>
<tr>
<td>( T_{in,turb} )</td>
<td>Inlet temperature to the turbine</td>
</tr>
<tr>
<td>( T_{pc} )</td>
<td>Temperature in the post combustor</td>
</tr>
<tr>
<td>( T_{com} )</td>
<td>Temperature in combustor</td>
</tr>
<tr>
<td>( \omega )</td>
<td>Turbine speed</td>
</tr>
<tr>
<td>( \dot{m}_{in,HX} )</td>
<td>Mass flow rate to heat exchanger</td>
</tr>
<tr>
<td>( \dot{m}_{in,pc} )</td>
<td>Mass flow rate to Post combustor</td>
</tr>
<tr>
<td>( \dot{m}_{CA} )</td>
<td>Mass flow rate cold air</td>
</tr>
<tr>
<td>( \dot{m}_{HA} )</td>
<td>Mass flow rate hot air</td>
</tr>
<tr>
<td>( \dot{m}_{BA} )</td>
<td>Mass flow rate bleed air</td>
</tr>
</tbody>
</table>

4.4.1 Gain Matrices

To obtain the gain matrices as defined in Eq. (3.3), the transfer function matrices must be obtained. Thus experimental design of inputs must be performed. Successful experiment design is critical to generating informative input/output data. Therefore a pseudo random binary sequence (PRBS) experimental input for multiple input multiple output system is designed using the guidelines
provided by (Gaikwad and Rivera, 1996). This guidelines aim for persistent excitation of input signal and statistical independence between the input and disturbances. The frequency range of interest \([\bar{\omega}_s, \bar{\omega}^*]\) is given by:

\[
\bar{\omega}_s = \frac{1}{\beta_s \tau_{dom}^H} \leq \bar{\omega} \leq \frac{\alpha_s}{\tau_{dom}^L} = \bar{\omega}^*
\]

(4.1)

where \(\alpha_s\) is the fractional closed loop speed of the response of the process, \(\beta_s\) is an integer representing the number of time constants that correspond to the settling time that is defined in this work as the time taken by the output(s) to reach and stay within 5% of the final value. The fastest (lowest) dominant time constant is represented by \(\tau_{dom}^H\) while the slowest (highest) dominant time constant is represented by \(\tau_{dom}^L\). To ensure excitation in the desired frequency range, the switching time of the PRBS is calculated to satisfy Eq. (4.2) (Gaikwad and Rivera, 1996). Eq. (4.3) is used to calculate the number of switches.

\[
T_{sw} \leq \frac{2.8\tau_{dom}^L}{\alpha_s}
\]

(4.2)

\[
N_s^{(1)} = 2^n_r - 1 \geq \frac{2\pi \beta_s \tau_{dom}^H}{T_{sw}}
\]

(4.3)

Here \(n_r\) is the number of shift registers and \(T_{sw}\) is the switching time. The PRBS sequence is repeated after \(N_s T_{sw}\) time units. The parameters \(\alpha_s\) and \(\beta_s\) are chosen to be 2.0 and 3.0 respectively.

The PRBS is designed with the following parameters as estimated from open loop tests. \(\tau_{dom}^L = 50\) s and \(\tau_{dom}^H = 150\) s. Fig. 4.6 shows the power spectrum of the PRBS of the bleed air valve signal for the time interval of [3000, 3500] (for clarity). The frequency is normalized to a range of \([0, \pi]\).

The plant used in the system identification is obtained from a first principles model developed by (Tsai et al., 2010). To identify linear time invariant models, two distinct simulations were run. Each of them lasting for 4500s. To begin, several parameters are loaded into the MATLAB® workspace which includes parameters for running the simulation, this includes initial conditions of the facility, reference parameters, piping parameters, pressure loss coefficients, air plenum physical data, compressor parameters and correlation/dimensionless numbers of the system. These are outlined in Table 2.13.
Figure 4.6 Power spectra of the PRBS for the experimental setup (top left) and the PRBS for bleed air valve (top right); Power spectra of the PRBS for the validation setup (bottom left) and the PRBS for bleed air valve (bottom right)
Table 4.4 Parameters loaded into the SIMULINK workspace

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ambient and reference conditions</th>
<th>Initial conditions</th>
<th>Piping/equipment parameters</th>
<th>Minor pressure loss coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinematic viscosity</td>
<td>Post combustor density</td>
<td>Internal diameters of pipes</td>
<td>Relative roughness</td>
<td></td>
</tr>
<tr>
<td>Dynamic viscosity</td>
<td>Pressure</td>
<td>Temperature post combustor</td>
<td>External diameters</td>
<td>Equivalent length</td>
</tr>
<tr>
<td>Air conductivity</td>
<td></td>
<td>Temperature post combustor surface</td>
<td>Insulation diameters</td>
<td>K factor</td>
</tr>
<tr>
<td>Compressor parameters</td>
<td>Reference temperature</td>
<td>Temperature combustor</td>
<td>Lengths</td>
<td></td>
</tr>
<tr>
<td>Prandtl number</td>
<td>Reference pressure</td>
<td>Temperature plenum</td>
<td>Air plenum parameters</td>
<td></td>
</tr>
<tr>
<td>Metal conductivity</td>
<td>Compressor map parameters</td>
<td>Temperature plenum surface</td>
<td>Heat exchanger parameters</td>
<td></td>
</tr>
<tr>
<td>Metal density</td>
<td>Plenum density</td>
<td>Turbine parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal Constant</td>
<td>Gas</td>
<td>Temperature compressor</td>
<td>Combustor parameters</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Inlet temperature to heat exchanger</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The nominal values of the steady state operating point was obtained from (Tsai et al., 2010) as 13.5g/s~39% ± 10%, 45kW±5kW, 14±4%, 40±10%, 40±10% for the fuel valve, load bank, bleed air, cold air and hot air bypass respectively. The gain matrices are obtained using the data
obtained from the simulation of the available SIMULINK file. The simulation is run with the fifth order accurate variable time step explicit ODE numerical solver: Dormand-Prince. Firstly, data is obtained from the first run to obtain results from which the transfer function is estimated. Secondly, data is generated with a distinct set of profiles for the manipulated variables and the disturbances to generate validation data for the estimated transfer functions. This process allows for model selection from the estimated transfer functions. The Akaike Final prediction error criterion was used for model selection as shown in Eq. (4.4):

\[
FPE = \text{det} \left( \frac{1}{N} \sum_{1}^{N} \epsilon(t, \hat{\theta}_N)\epsilon(t, \hat{\theta}_N)^T \right) \left( 1 + \frac{n_\theta}{N} \right) \left( 1 - \frac{n_\theta}{N} \right)
\]  

(4.4)

In Eq. (4.4), \(N\) is the number of values in the estimation data set, \(\epsilon(t)\) is the vector or prediction errors, \(n_\theta\) is the number of estimated parameters and \(\hat{\theta}_N\) is the vector of estimated parameters. The model classes differed in poles, zeros and time delays as can be utilized using “tfest” function from MATLAB®. These model classes from which selection was performed were restricted to pseudo first and pseudo second order transfer function models, i.e. the maximum difference between the number of poles and zeros is two. Figs. 4.7-4.8 show the comparison between the process (the Simulink model) and the model (the transfer function model) for the validation data set for two outputs (Temperature to plenum and temperature to turbine).
Figure 4.7 Comparison of model response (solid black) and process data (star blue) for Temperature to plenum.

Figure 4.8 Comparison of model response (dash dot black) and process data (star blue) for Temperature to turbine.
4.4.2 Estimation of Implementation Error

The implementation error captured by the diagonal matrix $W_n$ in Eq. (3.29) can be due to the measurement noise and other uncertainties (Kariwala et al., 2008). For the HyPer facility, it was assumed that the implementation error would solely stem forth from the measurement data. The experimental data from the HyPer facility is used to estimate the noise. Suppose the true value of a measured variable denoted by $\hat{y} \in \mathbb{R}^{n_y}$. Then the measured data $y \in \mathbb{R}^{n_y}$ is given by:

$$y = \hat{y} + \epsilon$$

(4.5)

To evaluate the magnitude of variance $\mathbb{E}[\epsilon^T \epsilon]$, the underlying estimate $\hat{y}$ must be estimated. For this problem, it is assumed that the noise is Gaussian and the variance of the noise is estimated by fitting the data with a discretized smoothing spline in (Garcia, 2010), the variance is then estimated from the corresponding residuals $\epsilon$. The smoothening of the data comes from the minimization of the residual sum of squares and a penalty $P(\hat{y})$ as given in Eq. (4.6). The degree of smoothing is controlled by the parameter $s$. The penalty is given as the tridiagonal matrix $D$ which is the second order difference matrix.

$$F(\hat{y}) = RSS + s P(\hat{y}) = \|y - \hat{y}\|^2 + s \|D\hat{y}\|^2$$

(4.6)

Minimizing Eq. (4.6) with respect to $\hat{y}$ yields

$$(I_n + s D^T D) \hat{y} = H^{-1} \hat{y} = y$$

(4.7)

The parameter $s$ is chosen to minimize the generalized cross validation score as proposed by (Craven and Wahba, 1978), this is given by:

$$s = \text{argmin} \ GCV \equiv \frac{RSS/n}{1 - tr(H)/n^2}$$

(4.8)

Where $RSS$ is the residual sum of squares given by $\|y - \hat{y}\|^2$ in Eq. (4.6) and $GCV$ is the generalized cross validation. Trace is denoted by $tr$. The number of samples is $n$. The estimated data $\hat{y}$ is obtained using discrete cosine transform (DCT) thus the noise variance is obtained as:

$$\mathbb{E}[(y - \hat{y})^T(y - \hat{y})] = \mathbb{E}[\epsilon^T \epsilon] = n \sum_i \left(\frac{1}{1 + s\lambda_i^2} - 1\right)^2 \text{DCT}^2_i(y)$$

(4.9)

Where $\lambda_i \forall i = 1, ..., n$ are the eigenvalues of the matrix $\Lambda$ obtained from the eigen-decomposition of $D$ as follows $D = U \Lambda U^{-1}$. This algorithm is applied to the experimental data from the HyPer facility. Extensive details are provided in (Garcia, 2010). Estimated noise variances are shown in
Table 4.5 while comparison of smoothed data and raw data of some of the measured variables is shown in Fig. 4.9.

**Figure 4.9** Data reconciliation showing smooth data obtained from noisy data.
Table 4.5 Estimates of noise variance for candidate controlled variable.

<table>
<thead>
<tr>
<th>S/n</th>
<th>Candidate controlled variable</th>
<th>Description</th>
<th>Noise variance</th>
<th>Std. deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>$\dot{m}_{in,pt}$</td>
<td>Mass flow rate to the plenum</td>
<td>3.9652E-10</td>
<td>1.99E-05</td>
</tr>
<tr>
<td>2.</td>
<td>$T_{in,pt}$</td>
<td>Temperature to the plenum</td>
<td>0.00051549</td>
<td>0.022704</td>
</tr>
<tr>
<td>3.</td>
<td>$T_{pl}$</td>
<td>Temperature in the plenum</td>
<td>0.0031992</td>
<td>0.056561</td>
</tr>
<tr>
<td>4.</td>
<td>$T_{in,turb}$</td>
<td>Inlet temperature to the turbine</td>
<td>0.0060983</td>
<td>0.078092</td>
</tr>
<tr>
<td>5.</td>
<td>$T_{pc}$</td>
<td>Temperature in the post combustor</td>
<td>0.0072492</td>
<td>0.085142</td>
</tr>
<tr>
<td>6.</td>
<td>$T_{com}$</td>
<td>Temperature in combustor</td>
<td>0.0072492</td>
<td>0.085142</td>
</tr>
<tr>
<td>7.</td>
<td>$\omega$</td>
<td>Turbine speed</td>
<td>10000</td>
<td>100</td>
</tr>
<tr>
<td>8.</td>
<td>$\dot{m}_{in,HX}$</td>
<td>Mass flow rate to heat exchanger</td>
<td>3.9652E-10</td>
<td>1.99E-05</td>
</tr>
<tr>
<td>9.</td>
<td>$\dot{m}_{in,pc}$</td>
<td>Mass flow rate to Post combustor</td>
<td>3.9652E-10</td>
<td>1.99E-05</td>
</tr>
</tbody>
</table>

4.4.3 Cost Function

The economic cost function for the HyPer facility is represented by the cost of producing electricity discounted by the profit of selling power to the grid. This is obtained according to the following procedure, first the compressor work and the losses is accounted for. These are given in kJ/s and are then converted into an equivalent cost in dollars. Similarly, the electricity drawn from the HyPer configuration is converted into an equivalent cost in dollars, these are implemented through the price of electricity. No cost is taken for the inflow of air to the HyPer facility, similarly, no cost is taken for the exhaust flue gas from the turbine. The cost of electricity for 2016 is given as 10.07cents/kWh. Next, the fuel flow is converted into dollars. Both the price of electricity and price of natural gas are obtained from US Energy Information Administration (EIA) [http://www.eia.gov/electricity/monthly/pdf/epm.pdf](http://www.eia.gov/electricity/monthly/pdf/epm.pdf). The cost of natural gas is obtained to be $2.45/MMBtu which translates to $0.84cents/kWh. Therefore the following is obtained:

\[-J\left(\frac{\$}{h}\right) = 0.1007(\dot{W}_{comp} + \dot{W}_{loss} - \dot{W}_{elec,turb} - \dot{W}_{elec,FC}) + 0.0084\dot{Q}\]  \hspace{1cm} (4.10)

Now as seen in Eq. (3.38), the second order derivatives of the cost function with respect to input and with respect to input and disturbance are needed. The cost function in Eq. (4.10) is not an explicit function of the input ‘u’ and disturbances ‘d’ thus the cost is evaluated from the data obtained from the process and this cost is regressed to a second order quadratic function in the
input space (see Table 4.2 for manipulated variables ‘u’ and disturbances ‘d’). Therefore, the parameters of the cost function to be employed in the controlled variable selection $J_{uu}, J_{ud}$ are determined. Fig. 4.10 shows a comparison of the regressed cost function with the objective function data from the process.

![Figure 4.10](image)

**Figure 4.10** Estimate of the scaled cost function (dash dot) and the process cost function (solid red)

### 4.5 Selection of Pareto sets with Multiagent Optimization

Again as in section 3.6.3.2, the optimization as defined in Eq. (3.45) was implemented on an Intel® Xeon® CPU E-5-1620 v2 with 32GB RAM using the heterogeneous multiagent framework programmed in MATLAB (Bankole et al., 2018b). A similar version with worst case loss defined in Eq. 3.26 was implemented in (Bankole et al., 2018c). The multiagent framework with parameters as described in Table 3.4.

The multiagent optimization takes approximately 57 seconds per solution totaling 15 minutes (while branch and bound optimization executes with a runtime of 40 minutes). The results obtained from the multiagent optimization must now be further analyzed in the posteriori analysis as discussed. The first consideration is the dependency of the controlled variables. As can be seen
from the Figs 4.11a - d. The cross-correlation function is used to obtain the similarity between the signals from the process. The cross correlation function for discrete signals $f$ and $g$ is defined as follows:

$$(f \ast g)[n] = \sum_{m=-\infty}^{\infty} f^*[n]g[m + n]$$ \hspace{1cm} (4.11)

Where the subscript $n$ is denoted as the lag. The following results are obtained:

**Figure 4.11** Cross correlation function for (a) $[T_{in,turb}, T_{com}]$ (b) $[T_{in,turb}, T_{pc}]$ (c) $[T_{com}, T_{pc}]$ (d) $[T_{pl}, T_{in,pl}]$

The cross correlation function in Fig. 4.11a through c show peaks close to zero lag which implies the correlation of variables $T_{in,turb}, T_{com}, T_{pc}$. Similarly, Fig. 4.11 (d) show peaks close to zero lag.
which implies the variables $T_{pl}$, $T_{in,pl}$ are correlated. All mass flow rates in the system are also correlated thus the controlled variable set reduces to that shown in Table 4.6.

**Table 4.6** Controlled variable Pareto set in descending order of optimality

<table>
<thead>
<tr>
<th>Controlled variable Set</th>
<th>Controlled variable</th>
<th>Econ($/h)</th>
<th>Controllability ($\sigma$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>1,2,5</td>
<td>34.75</td>
<td>1.00</td>
</tr>
<tr>
<td>C1</td>
<td>2,5,8</td>
<td>35.65</td>
<td>0.99</td>
</tr>
<tr>
<td>C2</td>
<td>1,3,5</td>
<td>36.26</td>
<td>0.91</td>
</tr>
<tr>
<td>C3</td>
<td>3,5,8</td>
<td>37.22</td>
<td>0.90</td>
</tr>
<tr>
<td>C4</td>
<td>2,5,9</td>
<td>31.46</td>
<td>0.76</td>
</tr>
<tr>
<td>C5</td>
<td>3,5,9</td>
<td>33.97</td>
<td>0.75</td>
</tr>
<tr>
<td>C6</td>
<td>3,5,7</td>
<td>81.33</td>
<td>0.77</td>
</tr>
<tr>
<td>C7</td>
<td>2,5,7</td>
<td>85.78</td>
<td>0.28</td>
</tr>
<tr>
<td>C8</td>
<td>1,3,6</td>
<td>31.61</td>
<td>0.76</td>
</tr>
<tr>
<td>C9</td>
<td>1,2,6</td>
<td>34.64</td>
<td>0.28</td>
</tr>
<tr>
<td>C10</td>
<td>3,6,8</td>
<td>37.31</td>
<td>0.25</td>
</tr>
<tr>
<td>C11</td>
<td>2,6,8</td>
<td>40.79</td>
<td>0.25</td>
</tr>
<tr>
<td>C12</td>
<td>3,6,9</td>
<td>74.44</td>
<td>0.32</td>
</tr>
<tr>
<td>C13</td>
<td>2,6,9</td>
<td>75.72</td>
<td>0.32</td>
</tr>
<tr>
<td>C14</td>
<td>1,5,7</td>
<td>162.71</td>
<td>0.38</td>
</tr>
<tr>
<td>C15</td>
<td>5,7,8</td>
<td>197.11</td>
<td>0.37</td>
</tr>
</tbody>
</table>

The top 16 results from the multiobjective optimization are shown in Table 4.6. Based on these results, it can be concluded that the controlled variables with the most self-optimizing performance are the mass flow rate to the plenum (1), the temperature in the plenum (2), the temperature in the post-combustor (5) i.e. set C0:[1,2,5]. This is because of minimal expected value of the economic loss ($34.75/h) and a high minimum singular value compared to other controlled variable sets. Therefore it exhibits the best compromise of economics and controllability at the nominal conditions. A Pareto plot of all controlled variable sets is given in Fig. 4.12. The sets at the top of the table are represented in lower right corner of Fig. 4.12, they represent lower economic loss and higher controllability. Contrarily, controlled variable sets at the bottom of Table 4.6 are depicted towards the left portion of Fig. 4.12. It should be noted that some controlled variable sets such as C8-C11 which offer great economic incentive are poor with controllability and as such cannot be selected.
It can be seen that all sets have some form of mass flow rate control. The control of the mass flow is important in this facility due to the coupled nature of the hyper facility. Transient disturbances in the mass flow rate can propel the system towards instability which leads to compressor surge and stall. This necessitates control of mass flow within the Hyper facility. Secondly, control of temperature is crucial. The turbine and the fuel cell are coupled via the exit temperature of the flue gas from the turbine, therefore the control of temperature especially the post-combustor temperature is crucial as this drives the turbine speed. If the temperature from the post-combustor is high, this would lead to a high turbine speed which in turn drives the compressor at higher speed as they are connected by the same shaft. Consequently, this leads to an increased airflow to the fuel cell leading to an overcool. Alternatively, if the temperature to the fuel cell is rather high, this would shorten the fuel cell life span. Therefore fluctuations in temperature are undesirable as it leads to thermal stress on the fuel cell. (Tucker et al., 2005). This therefore imposes the need for energy sink and sources to offset such transients therefore the bypass valves are highly pertinent.

**Figure 4.12** Pareto plots of the controlled variable selection problem.
4.6 Posteriori Analysis

In this section, the top results of the Pareto sets are subjected to off design conditions. This is done by changing the values of the disturbances (electric load to the turbine and the fuel flow rate) from the preset nominal conditions, this ranges from 80% of the nominal value of disturbances to 120% in steps of 10%. Then the process is run till it achieves steady state and the gain matrices are once again identified. This process is repeated at multiple off design conditions and the defined controllability function $J_c(c)$ (inverse of the minimum singular value $\sigma$ of the scaled gain matrix $\hat{G}$) is evaluated for the CV sets from the Pareto list in Table 4.6 (see Eqs. 3.44-3.45). For brevity, only the three sets which perform best at off design conditions are shown in Fig 4.13. These are sets C0, C1, C7. Due to the inherent nonlinearity of the process, it can be seen in Fig. 4.13a-b that the minimum singular value is not monotonic as the disturbances vary from 80% of the nominal to 120%. It can be inferred that set C1 is the best CV set to be chosen. This set has the best compromise between economics and controllability as well as at off design conditions.

![Controllability measure for sets C0 (square), C1 (circle), C7 (star) at off design points by varying a) Fuel flow rates, b) Electric load.](image)

**Figure 4.13** Controllability measure for sets C0 (square), C1 (circle), C7 (star) at off design points by varying a) Fuel flow rates, b) Electric load.
4.7 Conclusions

In this chapter, a top down control structure design was performed on a cyber physical gas turbine – solid oxide fuel cell HyPer facility. This involves three stages: a priori stage, optimization stage and finally posteriori evaluation of the top performing CV sets. This establishes the set of controlled variables which minimize economic drift from optimality as disturbances propagate through the system and yet pose minimal compromise with respect to controllability. Several variables were prescreened off during the apriori stage and an optimization scheme was formulated for selecting controlled variables based on a multiobjective function. The candidate controlled variables were chosen such that they satisfied the self-optimizing properties required. The results show that the optimal set exhibit a tradeoff between the economic and controllability cost function as expected. Furthermore, a novel multiagent metaheuristic platform is employed in this work which is computationally efficient compared to traditional branch and bound method which is rather exhaustive. This is highly pertinent for fast enumeration of CV sets for a processes such as cyber physical systems. Additionally, the optimal controlled variables address the possibility of transients and instability in the HyPer facility. The enduring challenge is to design a feedback control system that satisfactorily controls these identified variables.
Chapter 5

5 Real time Optimization

5.1 Introduction

Due to increasing concentration of CO₂ in the atmosphere, several efforts have been made in recent years at developing protocols for reducing anthropogenic emissions of CO₂ (Weaver et al., 2007). Emitting about 1.9 billion metric ton of CO₂ annually from coal-fired power plants, the United States contributes 33% of total energy related CO₂ emissions, out of which 81% of CO₂ emissions is from electricity generation facilities (Lin et al., 2012). Thus strong incentives exist for capturing CO₂ emissions from power plants and for minimizing the corresponding energy required by the capture processes.

Renewable energies such as that obtained from wind or solar can be instrumental in reducing CO₂ gas emissions. However, in cases of high penetration of renewables to the grid, fossil-based power plants need to follow a highly fluctuating power demand due to intermittency of the renewables, uncertainty in their availability, and variability in the amount of produced power. As renewable energy sources become more integrated into distributed power generation, load tracking of electricity demand becomes necessary (Carrasco et al., 2006). If the power generation plant includes a CO₂ capture unit, optimal scheduling of CO₂ capture operations would also become essential.

In view of economic operation of energy plants, the need for strategies to respond to seasonal, diurnal, or even hourly changes in electricity load and price has been suggested in literature. For example, (Cohen et al., 2010) suggested that electric power output can be increased to meet higher electricity demand by turning off the CO₂ capture plant in peak hours. According to (Chalmers et al., 2009), if CO₂ trading price is included, bypassing CO₂ capture is valuable when the $/MWh electricity selling price is 2–3 times higher than the $/ton penalty for not capturing the CO₂. (Lin et al., 2012) examined variability in electricity loads using an 11-hour peak and off-peak cyclical period. In tracking electricity load, peak loads resulted in lower CO₂ capture while off-peak electricity load resulted in higher CO₂ capture. Optimal scheduling of CO₂ capture was undertaken by (Sahraei and Ricardez-Sandoval, 2014), where the authors considered the control of a post-combustion capture of CO₂ using monoethanolamine (MEA). An optimal sequence of set points
for CO₂ capture was obtained by minimizing energy consumption and CO₂ emission. Two scenarios, namely high electricity generation and low CO₂ emission were considered. From the results, the authors concluded that the interactions of energy factors, environmental constraints and controllability were responsible for the differences in the optimal sequence of set points in both scenarios.

There is a scarcity of work in the open literature in the area of optimal scheduling of CO₂ capture and power production by optimizing the plant economics under various scenarios of carbon tax in the face of dynamic changes in the electricity price and demand. Two approaches have been proposed in the literature to achieve optimal economic operation of a plant. One approach is to employ a hierarchical structure as discussed by (Skogestad, 2004, Skogestad, 2000). In this case, an optimization layer determines optimal set points for the supervisory layer while the supervisory layer is designed for optimal tracking of the set point trajectory in the face of constraints and disturbances. Another option is to employ economic model predictive control as discussed by (Omell and Chmielewski, 2013, Ellis et al., 2014), where the objective function of the lower level controller considers economic variables in its objective function. It should be noted that the scheduling problem specific to CO₂ capture processes as part of an energy generating plant involves fluctuations in electricity demand and prices that evolve over a shorter time scale (on a minute or hourly scale), while taking into account CO₂ credit or deficit that needs to be considered over a much longer time period (such as days or months). Therefore, our approach to the optimal scheduling problem is to use the hierarchical structure that naturally facilitates separation of time scales. Therefore the scheduler (also known as the real-time optimizer above the supervisory layer) only solves the proposed optimization problem at a time interval that is appropriate for economic variables. This time interval is much longer than the time interval at which the lower level supervisory control needs to be executed. This multiscale feature specific to the scheduling problem of energy plants integrated with CO₂ capture units has not yet been studied in the open literature.

For formulating the optimal scheduling problem, predictions of unknown energy prices and demand, and optimal set points are defined within a ‘base period’ (or base time). The ‘base period’ is defined as the compliance period in which energy companies would be examined by the legislative bodies to comply with the legislative CO₂ capture requirements and would be taxed or provided incentive accordingly based on the aggregate emissions during the base period. While
the actual base period for CO$_2$ capture is expected to depend on a specific region or a country, it is anticipated that the base period for CO$_2$ capture may span several months to a year or more. It should be noted that as the CO$_2$ capture requirement is anticipated to be satisfied over the entire base period, as opposed to a desired target at any point of time, any discrepancy in the capture during previous time instances must be accounted for in the future or vice versa. This is not the case for typical control problems where the past deviation in the control objective from the set point might be completely neglected. This uncommon aspect anticipated for CO$_2$ capture has hardly been studied in the exiting literature.

In addition to optimal scheduling, the following three aspects need to be considered for optimal economic operation of the energy plant with CO$_2$ capture: design of the supervisory layer and control structure selection for the supervisory and regulatory control layers. An optimal design of the supervisory layer is essential for tracking the changing set point from the scheduler layer satisfactorily. Supervisory control layer design has been an area of active research for several decades now (Mckay et al., 1997, Bakshi and Stephanopoulos, 1994, Richalet, 1993). Supervisory controller design can be considered to be two separate, yet connected, problems: structural design (such as: what input(s) should be connected to what output(s), and how should they be connected) and controller design (such as: what type of controller should be used and how to tune those controllers for performance and robustness). As processes are expected to be more agile while operating close to the constraints, decentralized controls may become inadequate, requiring the need for centralized controllers (Wolff et al., 2014). Even though feasibility of centralized controllers involving fairly high number of variables is being realized recently due to the advent of powerful hardware and software, it is still intractable to solve one, single centralized controller that includes all controlled and manipulated variables for a large-scale plant. Neither is this approach necessary since the relative improvement in control performance by solving an increasingly larger problem typically keeps diminishing. Therefore the trade-off between the increasing computational expenses vs. diminishing returns needs to be evaluated. The objectives here are to optimally select the number of centralized controller(s), if any, and then determine the set of controlled (output) and manipulated (input) variables to consider in each of them. The works concerning the supervisory controller layer design is provided in the appendix.
Furthermore, stability conditions for the scheduler with electricity production and CO₂ capture are proposed. Lyapunov stability has been reported by (Huang et al., 2011) for cyclic steady-state processes. The authors have invoked Lipschitz continuity and weak controllability assumptions on the stage cost function and the model equations, respectively, for a generic nonlinear state space model. A similar approach is also investigated for infinite horizon nonlinear MPC which introduces a discount factor to keep the objective function bounded. Similarly, (Diehl et al., 2011) showed that under certain assumptions, asymptotic stability of an economic model predictive controller may be guaranteed by considering the stage cost as a function of deviation variables. In both works, valid constraints for the whole base period or cycle time of a cyclic process are not considered. Thus while it is possible for defined constraints to be satisfied at every given instant, an overall constraint may be violated. This is pertinent to the current problem description where an overall constraint on carbon capture must be considered. An example of this would be maximum carbon capture towards the end of the base period due to inaccurate predictions of electricity prices and demand at the beginning of the base period. Conditions for Lyapunov stability are discussed in detail in Section 5.3.

Finally, the methodology developed in this chapter is applied to an acid gas removal (AGR) unit as part of an integrated gasification combined cycle (IGCC) power plant. This technology has been the subject of research for several years (Chen and Rubin, 2009). The IGCC technology promises an efficient use of coal and the reduction of carbon emissions using pre-combustion capture compared to conventional power plants using post-combustion capture (Bhattacharyya et al., 2011). In the AGR unit, CO₂ is removed by a physical solvent such as SELEXOL due to high partial pressure of CO₂ thereby reducing the penalty for CO₂ capture. In addition to these merits, IGCC plants can follow load dynamically responding to the real-time price of electricity (Omell and Chmielewski, 2013). Furthermore, IGCC plants can be readily modified for both chemical and power production improving the controllability of the process in the face of fluctuating power demand (Robinson and Luyben, 2010).

In summary, this chapter focuses on optimal scheduling of advanced energy plants with CO₂ capture where load tracking as well as carbon capture targets for a given base period are both considered in the framework of an economic objective function (shown in Fig. 5.1). This is
achieved by considering not only changing electricity demand but also its prices with due consideration of penalty/incentive for violating/exceeding carbon capture targets. Effects of three different carbon tax scenarios on optimal scheduling of CO₂ capture and power production are evaluated. These scenarios are: no incentive for carbon capture; no incentive for carbon capture beyond a mandatory requirement; and lastly the trading of carbon emission allowances. Optimal set points for the extent of CO₂ capture and electricity production rates from the scheduler are then passed on to the supervisory control layer (see appendix for supervisory control layer design). Contributions of this chapter include: unique formulation of economic optimization with CO₂ capture including carbon tax, inclusion of past errors in the formulation of the scheduling problem, incorporation of different carbon tax scenarios and the development of Lyapunov stability conditions for optimal scheduling of a power system with CO₂ capture.

Figure 5.1 Overall flowchart of the optimization formulation and controller design
5.2 Mathematical Formulation

As the objective considered here is to maximize the profitability of the energy plant under consideration, the economic variables relevant to the optimal operation of the plant need to be considered. In addition to the typical operating costs of plants with CO$_2$ capture, real-time price of the product (i.e. electricity) and the effect of impending legislation on CO$_2$ capture need to be taken into account. Three different scenarios are considered to account for the effect of CO$_2$ capture legislation.

Scenario 1

In this scenario, all carbon emissions are charged at a fixed tax rate. Thus there is no allowable emission limit, nor is there any opportunity to trade CO$_2$ emission allowances.

Scenario 2

In this scenario, there is a penalty on CO$_2$ emissions above an allowable limit during the base period. However no reward whatsoever exists for capturing more CO$_2$ beyond this set limit. An example of this is found in (Sahraei and Ricardez-Sandoval, 2014) where the authors have noted that the U.S. Environmental Protection Agency has recently established a new limit for CO$_2$ emission of power plants wherein a new coal-fired power plant would need to meet a limit of 1100 lb. of CO$_2$ per MWh of electricity. In Maryland, for example, the legislation required payments of carbon tax beyond a limit of a million ton per year (Lu et al., 2012). An equivalent of this also applies to Alberta where a $15/ton taxation is applied for emissions beyond 100,000 ton of greenhouse gas annually (David, 2008). The implication of this scenario is to incentivize companies to capture at least the carbon target set by the regulatory agencies.

Scenario 3

Under this scenario, the so-called ‘cap and trade’ policy is evaluated. If a plant exceeds its cap on CO$_2$ emissions set by the regulatory agency, then it needs to buy the permit from the federal/state/local agency(ies) and/or from (private or non-private) organizations that are willing to trade CO$_2$ emission allowances. Therefore, if a plant captures more CO$_2$ than required, it can trade with others.
5.2.1 Forecasting Model

As the economic optimization needs to be carried out over the entire base period for evaluation of regulatory compliance for CO₂ emissions, a forecasting model is needed to generate future predictions of electricity prices and demand. Autoregressive integrated moving average (ARIMA) and state space models have been used in the open literature as forecasting models for electricity price and demand (Taylor et al., 2006, Taylor, 2010, Gould et al., 2008). A generic prediction model is considered as shown in Eq. (1), where \( d_k \) and \( e_k \) denote disturbances and stochastic noise at time step \( k \):

\[
d_{k+1} = F(d_k) + e_k
\]  

(5.1)

5.2.2 Economic Optimization Formulation

The formulation of the economic optimization scheme performed by the scheduler at time instant ‘\( i \)’ in the periodically spaced time horizons is given by the following:

\[
\max_u V(d, u, y, \delta)
\]

where:

\[
V(d, u, y, \delta) = \sum_{h=1}^H \left[ \sum_{k=i+m_{h-1}+1}^{i+m_h} w_{h,k} \left( f(u_{1,k}, d_{2,k|i}) - p(u_k) \right) \right] - J(y, \delta)
\]

subject to:

\[
d_{k+1} = F(d_k) + e_k
\]  

(5.2)

\[
y_k = F(x_k, v_k) + \omega_k
\]

\[
\delta = \sum_{k=1}^{i-1} y_{1,k} (y_{2,k} + y_{3,k} + y_{4,k} - z_{cap,k})
\]

\[
\Delta u_{\min} \leq \Delta u \leq \Delta u_{\max}
\]

\[
u_{\min} \leq u \leq u_{\max}
\]

\[
u_{1,k} \leq d_{1,k}
\]

\[
d = \begin{bmatrix}
\text{Electricity Demand} \\
\text{Electricity Price}
\end{bmatrix}
\]
The objective function $V(d, u, y, \delta)$ consists of revenue generation from electricity production $f$, a cost function for electricity production and the cost of carbon capture which is calculated using the function $p(u_k)$, this includes pumping and compressor costs, solvent make up, chilling, reboiling etc. The last term $J$ is a penalty cost function for carbon emission. This optimization is performed over the time span from the current time step $i$ to the end of the base period. This time span is subdivided into $H$ number of horizons, indexed by the variable $h$, and the cumulative number of time steps from the current time step $i$ to the horizon $h$, denoted by $m_h$. The argument involves the function $f(u_{1,k}, d_{2,k|i})$ evaluated as the estimated revenue generated from producing electricity. It should be noted that the decision variable $u$ and the predicted disturbances $d$ at any time step $k$ denoted by $u_k, d_k$ are both two tuple, thus the pair $u_{1,k}, d_{2,k|i}$ denotes the electricity production rate and the electricity price, respectively. The weights $w_{h,k}$ are used to imply the relative confidence in the accuracy of these terms. The economic penalty (or reward, if applicable) of carbon emission (or capturing more than mandated) is denoted by the function $J(y, \delta)$. The carbon tax penalty depends on cumulative past errors ($\delta$) in achieving the target carbon capture, this comprises all past errors from the beginning of the base time to current time step $i$. The maximum allowable mole fraction of greenhouse gas is denoted as $z_{cap}$. Therefore $J(y, \delta)$ varies based on the specific carbon tax scenario.

In Eq. (5.2), the process model is represented by a discrete time difference equation as shown above. In this model, $x, y, v$ and $\omega$ represent process states, outputs, inputs and noise. The values of the manipulated variables $v$ of the lower level controllers are implicitly dependent on the results of the optimization of Eq. (5.2). The set points obtained from (2) are passed as references $r$ to an MPC where $v$ is obtained from an optimization problem outlined later in Eq. (5.20). The constraint $u_{1,k} \leq d_{1,k}$ ensures that the electricity production never exceeds the demand apportioned to the power plant. The outputs of concern in the optimization are the flow of fuel/syngas $F_{out}$ which directly impacts electricity production and mole fractions of greenhouse gases ($CO_2$, $CO$ and $CH_4$). After every base time, the optimization problem is reset, assuming that $CO_2$ credits or taxes cannot
be carried forward from one base period to another. One of the features of the optimization scheme as posed in Eq. (5.2) is its coupled multiscale nature. It should be noted that the electricity production rate affects the fuel flowrate which, in turn, acts as a disturbance to the AGR unit where the CO₂ is captured. As the three different scenarios outlined earlier differ based on carbon tax scenario, the formulation for \( J(y, \delta) \) for the three scenarios is written as discussed.

**Scenario 1**

In this case, all expected emissions are simply penalized by a carbon tax rate \( \gamma_i \). For generality, this is allowed to vary, hence the subscript \( i \). Considering the typical carbon bearing components present in the feed gas to the AGR unit, Eq. (5.3) can be used to represent the penalty function where syngas/fluegas flow is denoted by \( y_1 \), the mole fraction of chemical species are denoted by the outputs \( y_2, y_3, y_4 \) and the past errors are denoted by \( \delta \). For post-combustion CO₂ capture, there is hardly any CO or CH₄ in the flue gas. For pre-combustion CO₂ capture, the syngas would contain all of these species plus some minor concentration of other carbon bearing species such as COS, CS₂, etc.

\[
J(y, \delta) = \gamma_i \cdot \left( \delta + \sum_{h=1}^{H} \sum_{k=i+m_{h+1}}^{i+m_h} w_{h,k} y_{1,k|i} (y_{2,k|i} + y_{3,k|i} + y_{4,k|i}) \right) \quad (5.3)
\]

where

\[
\delta = \sum_{k=1}^{i-1} y_{1,k} (y_{2,k} + y_{3,k} + y_{4,k}) \quad (5.4)
\]

**Scenario 2**

In this scenario, the tax function only becomes active if emitted carbon exceeds the allowed cap \( z_{cap} \). This scenario is represented by Eqs. (5.5), (5.6) and (5.7). The past cumulative contributions to the carbon released from the beginning of the base time to the current time step (due to excess or less CO₂ capture) is denoted by \( \delta \). At current time step \( i \), this is the first summation term in Eq. (5.6) while the second sum denotes estimation of future contributions to carbon emissions from all horizons to the penalty term. The carbon tax only exists as outlined in Eq. (5.7) if the \( \epsilon \) term is positive. This scenario reduces to scenario 1 if \( z_{cap} = 0 \).
\[ \delta = \sum_{k=1}^{i-1} y_{1,k} \left( y_{2,k} + y_{3,k} + y_{4,k} - z_{cap,k} \right) \]  

(5.5)

\[ \epsilon = \delta + \sum_{h=1}^{H} \left[ \sum_{k=i+m_{h-1}+1}^{i+m_{h}} w_{h,k} y_{1,k|i} \left( y_{2,k|i} + y_{3,k|i} + y_{4,k|i} - z_{cap,k} \right) \right] \]  

(5.6)

\[ J(y, \delta) = \begin{cases} 
\gamma_{buy,i} \cdot \epsilon & \forall \epsilon \geq 0 \\
0 & \forall \epsilon < 0 
\end{cases} \]  

(5.7)

**Scenario 3**

This is represented by Eqs. (5.8), (5.9) and (5.10). In this scenario, if more greenhouse gases are released, carbon credits are bought to compensate, denoted by the penalization term \( \gamma_{buy,i} \) i.e. the cost of purchasing credits. On the other hand, if less greenhouse gases are emitted than mandated, then equivalent credits can be sold at the rate of \( \gamma_{sell,i} \). This scenario reduces to Scenario 2 if \( \gamma_{sell,i} = 0 \). As before, the \( \delta \) term denotes past errors in meeting up with designated carbon capture target.

\[ \delta = \sum_{k=1}^{i-1} y_{1,k} \left( y_{2,k} + y_{3,k} + y_{4,k} - z_{cap,k} \right) \]  

(5.8)

\[ \epsilon = \delta + \sum_{h=1}^{H} \left[ \sum_{k=i+m_{h-1}+1}^{i+m_{h}} w_{h,k} y_{1,k|i} \left( y_{2,k|i} + y_{3,k|i} + y_{4,k|i} - z_{cap,k} \right) \right] \]  

(5.9)

\[ J(y, \delta) = \begin{cases} 
\gamma_{buy,i} \cdot \epsilon & \forall \epsilon \geq 0 \\
-\gamma_{sell,i} \cdot \epsilon & \forall \epsilon < 0 
\end{cases} \]  

(5.10)

**5.3 Stability**

In this section, a Lyapunov stability analysis is presented. However, unlike the work of Huang et al. (Huang et al., 2011) which presents the Lyapunov stability analysis of a cyclic process, the power plant integrated with CO₂ capture is considerably different since it is neither cyclic nor there is any desired steady state. Furthermore, as the CO₂ capture target is specified for a base period, discrepancy in the past CO₂ capture within a given base period must be accounted for. In addition,
as load-following is considered, plant dynamics strongly depends on forecasting of demand and price of electricity. Therefore, a Lyapunov stability analysis of this interesting system is undertaken.

A general discrete function for prediction of disturbances $d$ given by Eq. (5.1) is assumed. At time step $k$, let the stage cost of the optimization problem in Eq. (5.2) be denoted by $l(d_k, u_k, y_k, \delta_k)$ where $u_k$ denotes manipulated variables to be used as set points for the supervisory control layer beneath the scheduler. While these are outputs (degrees of freedom) of the scheduler optimization, they are inputs to the supervisory control layer. it should be noted that $y$ here denotes computed values of future outputs as past outputs have been absorbed in $\delta$. It is proposed that the economic objective function, defined as in Eq. (5.2), is a Lyapunov function under certain assumptions.

**Assumption 1**
The underlying process defined by $y_k = \mathcal{F}(x_k, v_k)$ is controllable, i.e. For any final time $t > 0$ and any initial state $x_0$, there exists a control that transfers the state to the desired value at time $t$.

**Assumption 2**
The predictions denoted by $F(d_k)$ and the stage cost $l(d_k, u_k, y_k, \delta_k)$ are both Lipschitz continuous on the set of all $u \in \mathbb{U}, d \in \mathbb{D}, y \in \mathbb{Y}, \delta \in \mathbb{D}$ with Lipschitz constants $l_f, l_l \geq 0$ such that $\forall u \in \mathbb{U}, d \in \mathbb{D}, y \in \mathbb{Y}, \delta \in \mathbb{D}$ this gives the following:

$$|F(d_1) - F(d_2)| \leq l_f|d_1 - d_2| \quad (5.11)$$

$$|l(d_1, u_1, y_1, \delta_1) - l(d_2, u_2, y_2, \delta_2)| \leq l_l|(d_1, u_1, y_2, \delta_1) - (d_2, u_2, y_2, \delta_2)| \quad (5.12)$$

**Assumption 3**
In the absence of estimation errors of the forecasting model, let the optimal sequence of set points obtained from the optimization at time step $k$ of the base period be denoted by $\bar{u}^*(x, k) = \{u^*(k|k), u^*(k + 1|k), ..., u^*(N|k)\}$ and optimally computed outputs be denoted by $\{y^*(k|k), y^*(k + 1|k), ..., y^*(N|k)\}$. For the inputs $u$, predicted outputs $y$ and control errors $\delta$, it is assumed that there exists $\kappa_\infty$ functions $\beta(\cdot), \gamma(\cdot), \xi(\cdot)$ such that for some arbitrary time step $k = i$, such that

$$\sum_{k=i}^{N} |u(k|i) - u^*(k|i)| \leq \sum_{k=i}^{N} \beta(|d_k - d_k^*|) \quad (5.13)$$
\[
\sum_{k=i}^{N} |y_{k|i} - y_{k|i}^*| \leq \sum_{k=i}^{N} \gamma (|d_k - d_k^*|) \tag{5.14}
\]
\[
\sum_{k=i}^{N} |\delta_k| \leq \sum_{k=i}^{N} \xi (|d_k - d_k^*|) \tag{5.15}
\]

Where the predictions and the actual values of the disturbances are denoted by \(d\) and \(d^*\) respectively. The above assumption ensures that the deviation in scheduler outputs (degrees of freedom) and the lower level computed outputs remain bounded as the true value of the disturbances deviate from the optimal prediction.

**Assumption 4**

The optimization problem defined in Eq. (5.2) satisfies the linear independent constraint qualification (Nocedal and Wright, 2006), sufficient second order conditions (Nocedal and Wright, 2006) and strict complementarity (Nocedal and Wright, 2006) at the solution.

Assumption 4 indicates that Eq. (5.2) is well-posed and thus a solution exists which is locally unique. The formulation defined by Eq. (5.2) is transformed to the form of deviation variables from the optimal (Diehl et al., 2011, Huang et al., 2011). Thus

\[
\begin{align*}
\tilde{d}_k &= d_k - d_k^* \\
\tilde{u}_k &= u_k - u_k^* \\
\tilde{y}_k &= y_k - y_k^* \\
\tilde{\delta}_k &= \delta_k - \delta_k^*
\end{align*}
\tag{5.16}
\]

Thus the transformed disturbance model evolves according to the following:

\[
\tilde{d}_{k+1} = F(\tilde{d}_k + \tilde{d}_k^*) - d_{k+1}^* = \bar{F}(\tilde{d}_k)
\tag{5.17}
\]

where \(\bar{F}(0) = 0\)

For the transformed system, the stage cost is modified as follows:

\[
\bar{l}(\tilde{d}, \tilde{u}, \tilde{y}, \tilde{\delta}) \triangleq l(d, u, y, \delta) - l(d^*, u^*, y^*, \delta^*)
\tag{5.18}
\]

**Lemma 1**

The stability of the transformed formulation with stage cost \(\bar{l}(\tilde{d}, \tilde{u}, \tilde{y}, \tilde{\delta})\) at \(0,0,0,0\) is equivalent to the stability of the original system with stage cost \(l(d, u, y, \delta)\) at \((d^*, u^*, y^*, \delta^*)\).
Assumption 4 implies that a unique solution exists to the optimization problem formulated in Eq. (5.2) and \( \sum_{k=1}^{N} l(d^*, u^*, y^*, \delta^*) \) is a constant thus the solution to the optimization problem using Eq. (5.18) as a stage cost is the same as the solution obtained with Eq. (5.2).

Thus the objective function is similarly transformed

\[
\bar{V}(i) = \sum_{k=i}^{N} l(d_k, u_k, y_k, \delta_k)
\]

\[
\bar{V}(i) = \sum_{k=i}^{N} \left( l(d^*_k + d^*_k, u_k + u^*_k, y^*_k + y^*_k + \delta_k^* + \delta^*_k) - l(d^*_k, u^*_k, y^*_k, \delta^*_k) \right)
\]

From Eq. (5.20), it is evident that \( \bar{l}(0,0,0) = 0 \). Therefore the stability of the transformed formulation in Eq. (5.19) is equivalent to the stability of the original system in Eq. (5.2).

Due to the Lipschitz continuity of the prediction model and cost function, it is apparent that the transformed system and the cost function are Lipschitz continuous i.e. \( \forall d \in \mathbb{D}, u \in \mathbb{U} \) there exists Lipschitz constants \( \bar{l}_f \) and \( \bar{l}_l \) such that

\[
|\bar{F}(d_1) - \bar{F}(d_2)| \leq l_f |d_1 - d_2|
\]

\[
|l(d_1, u_1, y_1, \delta_1) - l(d_2, u_2, y_2, \delta_2)| \leq l_l |(d_1, u_1, y_1, \delta_1) - (d_2, u_2, y_2, \delta_2)|
\]

In addition, this implies that the transformed system has bounded inputs, future outputs and errors such that:

\[
\sum_{k=i}^{N} |\bar{u}(k|i)| \leq \bar{\beta} |\bar{d}_i - 0|
\]

\[
\sum_{k=i}^{N} |\bar{y}_k| \leq \bar{\psi}(\bar{d}_i - 0)
\]

\[
\sum_{k=i}^{N} |\bar{\delta}_k| \leq \bar{\xi}(\bar{d}_i - 0)
\]

**Assumption 5**

There exists a \( \kappa_\infty \) function \( \psi(\cdot) \) such that the stage cost \( \bar{l}(d_k, u_k, y_k, \delta_k) \) satisfies

\[
\bar{l}(d, u, y, \delta) \geq \psi(|d - 0|)
\]
**Theorem:**

Based on assumptions 1-5, then $V(i)$ as defined by Eq. (5.2) is a Lyapunov function and the transformed formulation defined by Eq. (5.19) is asymptotically stable at $(0,0,0,0)$.

**Proof:**

The following is obtained:

$$
\bar{V}(i + 1) - \bar{V}(i) = -\bar{I}(\bar{d}_i, \bar{u}_i, \bar{y}_i, \bar{\delta}_i) \leq -\psi(|\bar{d}_i - 0|) \quad (5.27)
$$

From assumption 5, Eq. (5.19) and Eq. (5.20), and triangle inequality, one obtains

$$
\bar{V}(i) = \sum_{k=i}^{N} \bar{I}(\bar{d}_k, \bar{u}_k, \bar{y}_k, \bar{\delta}_k)
$$

$$
= \sum_{k=i}^{N} \left( \bar{I}(\bar{d}_k, \bar{u}_k, \bar{y}_k, \bar{\delta}_k) - \bar{I}(0,0,0,0) \right) \quad (5.28)
$$

$$
\leq l_{\bar{I}} \left( \sum_{k=i}^{N} |\bar{d}_k - 0| + \sum_{k=i}^{N} |\bar{u}_k - 0| + \sum_{k=i}^{N} |\bar{y}_k - 0| + \sum_{k=i}^{N} |\bar{\delta}_k - 0| \right)
$$

From the Lipschitz continuity of $\bar{F}(\cdot, \cdot)$, one obtains

$$
|\bar{d}_k - 0| \leq l_{F}^{k-i} |\bar{d}_i - 0| \quad (5.29)
$$

$$
\sum_{k=i}^{N} |\bar{d}_k - 0| \leq L_F [ |\bar{d}_i - 0|] \quad (5.30)
$$

where $L_F \geq \sum_{k=i}^{N} l_{F}^{k-i}$

Substituting Eqs. (5.30), (5.23), (5.24) and (5.25) into Eq. (5.28), the following is obtained

$$
\bar{V}(i) \leq l_{\bar{I}} \left( L_F [ |\bar{d}_i - 0|] + \bar{\beta}( |\bar{d}_i - 0|) + \bar{\gamma}( |\bar{d}_i - 0|) + \bar{\xi}( |\bar{d}_i - 0|) \right) \quad (5.31)
$$

Therefore

$$
\bar{V}(i) \leq \Phi(|\bar{d}_i|) \quad (5.32)
$$

where $\Phi(\cdot) = l_{\bar{I}} \left( L_F (\cdot) + \bar{\beta}(\cdot) + \bar{\gamma}(\cdot) + \bar{\xi}(\cdot) \right)$ is a $\kappa_\infty$ function. The set of equations in (5.26), (5.27) and (5.32) completes the proof that the function $\bar{V}(\cdot)$ defined as in Eq. (5.19) is a Lyapunov
function. Therefore with respect to Lemma 1, the original system $V(\cdot)$ defined as in Eq. (5.2) is a Lyapunov function. ■

5.4 Case Study

The proposed optimal scheduling algorithm is implemented on an IGCC power plant with CO$_2$ capture based upon the model developed by (Bhattacharyya et al., 2011). In this process, the syngas, mainly hydrogen and carbon monoxide, produced in the gasifier is sent to a series of water gas shift reactors (modeled as adiabatic plug flow reactors in series) with inter-stage cooling. The shifted syngas is then sent to the AGR unit where CO$_2$ and hydrogen sulfide (H$_2$S) are selectively absorbed from the syngas leaving mostly hydrogen in the clean syngas. The cleaned syngas is then sent to the gas turbine for power production. The hot exhaust gas from the gas turbine is then sent to a heat recovery steam generator where it is used to raise steam at various pressures for additional power production. Readers are referred to (Bhattacharyya et al., 2011) for a comprehensive discussion.

5.4.1 Problem formulation

The terms in the cost function for electricity production and carbon capture in Eq. (5.2) need to be specified before one can proceed.

$$f(u_{1,k}, d_{2,k|i}) = E_{k|i} \cdot F_s \cdot \eta \cdot LHV$$ (5.33)

In Eq. (5.33), $f(u_{1,k}, d_{2,k|i}), E_{k|i}, F_s$, and $\eta$ denote the revenue, electricity price, syngas flow rate, and overall efficiency for converting the syngas to electricity defined based on the lower heating value (LHV). The cost of carbon capture $p(u_k)$ is obtained from the least squares regression to fit a quadratic model as shown in Eq. (5.34). As stated in Section 5.2.2, this cost includes pumping and compressor costs, solvent make up, chilling, reboiling etc. as can be seen in Section 3.1.1 of (Jones et al, 2014) and Eq. (3.57). To reduce computational complexity, two horizons namely the near and far horizon are used. In this case, the weights $w_h$ in Eq. (5.2) reduce to $w_1$ and $w_2$. The relative change in CO$_2$ capture due to change in the concentration of CO and CH$_4$ in the outgoing stream from the SELEXOL unit are neglected since partial pressure of these species at the SELEXOL unit inlet is low resulting in negligible capture in the absorbers.

$$p(u_k) = \left(\frac{50F_{s,k}z_{c,k}}{F_{s,\text{max}}} + 0.198z_{c,k}^2\right)/11.58$$ (5.34)
Since this study is focused on a single power generation unit while in real life a large number of power generators would participate in following the load, the electricity demand data available from the grid are scaled such that the IGCC plant considered here is able to provide the maximum electricity demand. The maximum syngas flow rate to the process is given as $2.79 \times 10^4$ kmol/h while minimum syngas flow is set at $4.53 \times 10^3$ kmol/h. Nominal flow is set at $1.63 \times 10^4$ kmol/h. Minimum carbon capture is set at $55\%$ while maximum carbon capture is set at $97\%$. Efficiency $\eta$ and the LHV of syngas are assumed constant and the product $\eta \cdot \text{LHV}$ is set as $154.3$ MJ/kmol. The target CO$_2$ capture for the second and third scenarios is set to be $80\%$. The carbon tax $\gamma$ is set at $100/\text{ton carbon}$ (Poterba, 1991). For simplicity, the base period is set to be three months while the near horizon is set to two weeks. The objective function differs for each scenario due to the difference in the carbon tax penalty term which is outlined for the different scenarios below. As the penalty term $J(y, \delta)$ is based on the CO$_2$ released, in the formulations that follow $z_{c,k}$ denotes the carbon capture setpoint from the scheduler at time step $k$, while $\alpha$ denotes the carbon capture target set by regulating agencies. It should be noted that the decision variable $u$ here consists of $[F_S, z_c]$.

**Scenario 1**

\[
V(d, u, y, \delta) = \sum_{k=1}^{i+m} w_{1,k} \left( E_{k|i} F_{S,k} \eta \text{LHV} - p(u_k) \right) \\
+ \sum_{k=i+m+1}^{N} w_{2,k} \left( E_{k|i} F_{S,k} \eta \text{LHV} - p(u_k) \right) - J(y, \delta) \\
J(y, \delta) = \gamma \left( \delta + \sum_{k=i}^{i+m} w_{1,k} F_{S,k|i} z_{CO_2,k}(1 - z_{c,k|i}) \right) \\
+ \sum_{k=i+m+1}^{N} w_{2,k} F_{S,k|i} z_{CO_2,k}(1 - z_{c,k|i}) \right) \\
\]

Where $\delta = \sum_{k=1}^{i} F_{S,k} z_{CO_2,k}(1 - z_{c,k})$

Eqs. (5.35) and (5.36) follows Eqs. (5.2), (5.33), (5.3) and (5.4), represent the objective function for Scenario 1. In Eq. (5.36), the product $F_S z_{CO_2}$ denotes the flow of CO$_2$ into the AGR unit and
$z_{c,k}$ denotes the fraction of CO$_2$ captured at time $k$ while $z_{c,k|i}$ denotes future CO$_2$ capture fraction set points from the scheduler.

**Scenario 2**

$$V(d, u, y, \delta) = \sum_{k=i}^{i+m} w_{1,k} \left( E_{k|i} F_{s,k} \eta \text{LHV} - p(u_k) \right) + \sum_{k=i+m+1}^{N} w_{2,k} \left( E_{k|i} F_{s,k} \eta \text{LHV} - p(u_k) \right) - J(y, \delta) \tag{5.37}$$

$$\epsilon = \delta + \sum_{k=i}^{i+m} w_{1,k} F_{s,k|i} z_{C_{O2},k} (\alpha - z_{c,k|i}) + \sum_{k=i+m+1}^{N} w_{2,k} F_{s,k|i} z_{C_{O2},k} (\alpha - z_{c,k|i}) \tag{5.38}$$

$$J(y, \delta) = \begin{cases} 
\gamma_{b_u} \cdot \epsilon & \forall \; \epsilon \geq 0 \\
0 & \forall \; \epsilon < 0 
\end{cases} \tag{5.39}$$

where $\delta = \sum_{k=1}^{i-1} F_{s,k} z_{C_{O2},k} (\alpha - z_{c,k})$.

Scenario 2 is represented by Eqs. (5.37) - (5.38). In Eq. (5.39), the discrepancy in carbon capture (i.e. difference between target and actual capture) in the past is denoted by the first term while the second and third terms denote predicted differences in CO$_2$ capture in the near and far horizon.

**Scenario 3**

$$V(d, u, y, \delta) = \sum_{k=i}^{i+m} w_{1,k} E_{k|i} \left( F_{s,k} \eta \text{LHV} - p(u_k) \right) + \sum_{k=i+m+1}^{N} w_{2,k} E_{k|i} \left( F_{s,k} \eta \text{LHV} - p(u_k) \right) - J(y, \delta) \tag{5.40}$$

$$\epsilon = \delta + \sum_{k=i}^{i+m} w_{1,k} F_{s,k|i} z_{C_{O2},k} (\alpha - z_{c,k|i}) + \sum_{k=i+m+1}^{N} w_{2,k} F_{s,k|i} z_{C_{O2},k} (\alpha - z_{c,k|i}) \tag{5.41}$$

$$J(y, \delta) = \begin{cases} 
\gamma_{b_u} \cdot \epsilon & \forall \; \epsilon \geq 0 \\
-\gamma_{s_e} \cdot \epsilon & \forall \; \epsilon < 0 
\end{cases} \tag{5.42}$$

Eqs. (5.40-5.42) denote the objective function for Scenario 3. Here, $\delta$ is computed as before in Scenario 2.
5.4.2 Forecasting model

Here, a stochastic state space model shown in Eq. (5.43) has been identified using historical hourly electric price and grid load for the year 2014 for the Texas grid obtained from www.pjm.com.

\[
\begin{align*}
q_{k+1} &= \bar{A}q_k + \bar{B}e_k \\
d_{k+1} &= \bar{C}q_k
\end{align*}
\]

Forecasting models for the electricity price and demand are of the form given by Eq. (5.43). In the near horizon, forecasting is expected to be more accurate than the longer range predictions. In addition, monthly predictions are arrived at by scaling with respect to the standard deviation for each month to accurately simulate the monthly variations in prices and demand. Figs. 5.2 and 5.3 show the goodness of fit for the forecasting model. It can be observed that the model predicts the daily change in electricity demand and prices within given tolerance. During the optimization at any time step \(k\) when the scheduler runs, the discrete time model employs an error term calculated as the difference between previous forecasts for the current time step and actual values, which is then used to adjust future predictions. This accounts for unexpected variations in load and price of electricity. The forecasting model is then used to model electricity demand and price data obtained from www.pjm.com for the Commonwealth Edison utility company serving Illinois area. The purpose of using a different data set than those that were developed to simulate the forecasting model is to simulate errors in the forecasting model and study performance of the scheduler in the face of such errors.
Figure 5.2 Comparison of predicted and actual scaled electricity prices (data from www.pjm.com for the Texas grid for 2014).

Figure 5.3 Comparison of predicted and actual scaled electricity demand (data from www.pjm.com for the Texas grid for 2014).
5.4.3 Software Implementation

The MPC control toolbox in MATLAB Simulink was used to develop the controller and the Simulink flowsheet was coupled with MATLAB workspace where the scheduler is being run through interfacing blocks within Simulink.

At any time step, when the scheduler runs, the present price and demand, current run step and past values of the controlled variable and other variables are fed into the MATLAB script that calls the function which performs the optimization. Therein the prices and demand are scaled depending on the present month determined based on the current time. Next, future predictions are made using an identified stochastic model. Then the optimizer employs the sequential quadratic programming method (SQP) implemented in MATLAB’s ‘fmincon’ function. Once the optimal setpoints are obtained, the current set point is passed to the MPC and the process unit in Simulink where the discrete state space linear model is used to model the AGR unit of the IGCC plant. There, a variable step size ODE solver (Dormand-Prince) is employed to simulate the process from the current time step to the next time step (i.e. one hour). Actual values of the CVs are stored in the MATLAB workspace and are transferred back to the workspace of the script which calls the optimization function. The process is repeated until the end of the base time. The sequence of steps and interaction between the MATLAB and Simulink blocks is shown in Fig. 5.4 while the simulation flowsheet as implemented in Simulink is shown in Fig. 5.5.
Figure 5.4 Flowsheet of scheduler control problem

Figure 5.5 Schematic of the simulation flowsheet in Simulink
5.4.4 Results

In computing the optimal sequence of syngas flow rates and carbon capture rate, the data on electricity price and demand were obtained from http://www.pjm.com/markets-and-operations/data-dictionary.aspx for January through March 2015 for the ComEd utility company and were used for all three scenarios. Real time electricity demand and prices are shown in Figs. 5.6-5.7. To render the results and the data more amenable to visual analysis, the electricity prices and demand were averaged out on a daily basis thus reducing data points from 2160 (hour) to 90 (day) as shown in Figs. 5.9a and 5.9b. The proposed algorithm can work on longer periods but 90 days is chosen here for the sake of brevity. Prediction errors of the developed forecasting model on both electricity prices and demand are shown in Fig. 5.8a and 5.8b respectively.

![Electricity demand variation](image)

**Figure 5.6** Hourly variation of electricity demand from January to March, 2015 for the ComEd utility company
Figure 5.7 Hourly variation of electricity price from January to March, 2015

Figure 5.8 Errors in predictions of (a) electricity demand and (b) price from January to March, 2015.
Figures 5.9 Daily averaged electricity price (a) and demand (b) from January to March, 2015

Figs. 5.10-5.12 show the change in the syngas flowrate and CO₂ capture rate for Scenarios 1-3, respectively. For all three scenarios, the syngas flow rate is correlated with the electricity demand and one can observe a general trend in the sequence of Figs. 5.10a, 5.11a, and 5.12a.

For Scenario 1 in Fig. 5.10, tax is levied on all carbon released so the optimal sequence obtained from the RTO (Real Time Optimizer) keeps the CO₂ capture rate at its maximum. This occurs as long as the tax levied on the CO₂ released is greater than the cost of CO₂ capture. Contrarily, if no RTO was in place, this would be at a nominal carbon capture set point. No interesting interplay exists between energy demand, electricity prices and the optimal set point policy during operation. However, as shown later, as the carbon tax is lowered, the cost of carbon capture becomes higher than the carbon tax penalty levied and thus Scenario 1 can result in lower CO₂ capture than the nominal target.
Scenario 1

**Figure 5.10** Syngas flow rates (a) and CO₂ capture fraction set points (b) for the Scenario 1

For Scenarios 2 and 3, when electricity prices are low, higher incentives exist to capture higher percentages of CO₂ and conversely, when the electricity prices peak, the scheduler decreases the carbon capture set points accordingly. Dips in the electricity prices can be seen to correspond to peaks in the CO₂ carbon capture. Two prominent peaks can be seen around day 50, and the resulting troughs in CO₂ capture can be seen in both Scenarios 2 and 3 as shown in Figs. 5.11b and 5.12b. The results of Scenario 3 are similar to Scenario 2 with respect to higher CO₂ capture targets at low electricity prices and vice versa, however in contrast to Scenario 2, when the buying and selling price of CO₂ credit are both high, and the selling price is higher than the cost of CO₂ capture, maximum CO₂ capture is generally preferred by the optimizer in order to sell prospective CO₂ credit. In contrast, when the selling price is very low and the buying price is equally low, and lower than the cost of CO₂ capture, the scheduler sets minimum CO₂ capture targets.
Scenario 2

![Figure 5.11](a) Syngas flow rates and (b) CO₂ capture fraction set points for Scenario 2

Scenario 3

![Figure 5.12](a) Syngas flow rates and (b) CO₂ capture fraction set points for Scenario 3

The relative advantage of the RTO with respect to plant operation at nominal conditions can be seen in Figs. 5.13-5.15 for Scenarios 1-3, respectively. For each scenario, a plot of the total profit objective function value and a plot of the operational cost of carbon capture are placed side by
side. For Scenario 1, higher values of revenue can be generated by exploiting the stochastic predictions of electricity demand from the grid as opposed to nominal power production (Fig. 5.13a). This exploitation however leads to an increase in CO$_2$ capture cost compared to the nominal case as seen in Fig. 5.13b. For Scenario 2, similar arguments apply to the total objective function as shown in Fig. 5.14a, however the operational cost of the plant is lower than the nominal case. This is due to the optimized carbon capture profile where higher percentages of carbon capture are scheduled for periods with lower electricity prices and vice versa. Therefore while the nominal case sets a constant power production and carbon capture, exploitation of the electricity prices enables the scheduler to achieve lower operational cost for the AGR unit (Fig. 5.14b). Close inspection shows that the RTO achieves higher costs with respect to the nominal case between day 1 and day 20, however this is only due to the higher carbon capture at the beginning of the base time as electricity prices are relatively low (cf Fig. 5.9b).

Finally for Scenario 3 in Fig. 5.12, due to the flexibility of trading tax credits, the difference in the profit objective function value with respect to the nominal case is significant. Higher values of revenue in the beginning of the base period can be seen due to higher carbon capture as the real time optimizer takes advantage of lower electricity prices in the beginning of the base period (cf Fig. 5.9b) while capturing relatively lower amounts of carbon emissions towards the end of the base period. Similar to Scenario 2, this corresponds to a higher carbon capture cost as shown in Fig. 5.16b. In this scenario, both buying and selling prices of CO$_2$ are set at $100/ton (see section 5.4.1), thus the high selling price motivates the scheduler to capture an average of 90% of the overall carbon (cf fig 15b) as compared to the nominal target of 80% as specified in Section 5.4.1. This difference renders prospective selling units of CO$_2$ credits available for revenue. Thus the marked difference in the objective function for the RTO and nominal case in Fig. 5.15a. This corresponds to a significant increase in the cost of carbon capture for the AGR unit. For a 90 day period, the overall values are summarized in Table 5.1.

**Table 5.1** Summary of objective function values and cost of carbon capture for all scenarios

<table>
<thead>
<tr>
<th></th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RTO</td>
<td>No RTO</td>
<td>RTO</td>
</tr>
<tr>
<td>Objective function ($)× 10^7</td>
<td>3.82</td>
<td>3.38</td>
<td>6.33</td>
</tr>
<tr>
<td>CO$_2$ capture cost ($)× 10^5</td>
<td>2.22</td>
<td>2.00</td>
<td>1.73</td>
</tr>
</tbody>
</table>
**Figure 5.13** Objective function (Profit) and cost of carbon capture for Scenario 1 (with RTO – black solid, without the RTO – blue dash dot).

**Figure 5.14** Objective function (Profit) and cost of carbon capture for Scenario 2 (with RTO – black solid, without the RTO – blue dash dot).
Results from the RTO strongly depend on the amount of tax levied on carbon emissions. Therefore, a study is conducted to evaluate the RTO dynamics due to change in the carbon tax ($/ton). For Scenarios 1 and 2, the tax levied on the \( \text{CO}_2 \) emissions was varied and for Scenario 3, the geometric mean of the buying and selling credit for \( \text{CO}_2 \) emissions was varied. These tax values were varied until the minimum and maximum carbon capture is reached for each scenario. The average carbon capture during the entire base period is plotted against different tax values as shown in Fig 5.16.
Figure 5.16 Sensitivity of average carbon capture due to changes in the carbon tax $\gamma$($/\text{ton CO}_2$)
(a) Scenario 1; (b) Scenario 2; (c) Scenario 3.

From Figs. 5.16a-c, it is apparent that reduced tax levied on CO$_2$ emissions results in a concomitant
decrease in the amount of CO$_2$ captured as expected. In the case of Scenario 2, the maximum CO$_2$
captured is the target specified for the base time as no credit is gained by capturing more CO$_2$. For
Scenario 3, a similar trend is observed where the CO$_2$ capture varies from the maximum to
minimum as the geometric mean of the buying and selling price of CO$_2$ decreases. However,
greater incentives to capture higher carbon exists for Scenario 3 due to the possibility of selling
CO$_2$ credits hence the sharp rise in the CO$_2$ capture fraction at $\gamma_{buy} = \gamma_{sell} = 0.08$/ton CO$_2$.

5.5 Conclusions

Optimal scheduling of an IGCC power plant with CO$_2$ capture is provided in this study. A complete
mathematical formulation including revenue generation and operational cost of carbon capture
under different tax scenario is presented. In addition to this, Lyapunov-based stability conditions
are provided for which the results are guaranteed for the optimizer. Effects of three different
scenarios for carbon tax on optimal set points of syngas flowrate and CO$_2$ capture are investigated.
As expected, carbon capture for all scenarios considered is negatively correlated with electricity
price. Results show how exploitation of the stochastic predictions of electricity price and demand
can result in increased profits for power plants. For Scenario 1, it is shown that higher values of
profit can be obtained by producing just the required amount of electricity to offset power produced by fluctuating sources such as renewables. For Scenario 2, in addition to higher values of profit which is obtained in part by load following, reduced cost of carbon capture is obtained by exploiting variation in electricity prices. Lastly for Scenario 3, higher values of profit are obtained due to three properties, one is the exploitation of electricity demand, secondly the electricity prices are used to schedule carbon capture, lastly prospective selling or buying of CO$_2$ credits can be taken advantage of to arrive at optimal scheduling of power production and carbon capture. Of all scenarios, Scenario 3 takes most advantage of the scheduling as can be seen in the sensitivity of the percentage of CO$_2$ capture to changes in tax prices. The insights gained from this study can be applicable to real power plants for increasing profit and revenue without violating environmental constraints of carbon capture especially when the operational cost of running the plant is high.
Chapter 6

6 Recommendations and future research

The contributions of this thesis include the development of a connectivity estimation procedure and algorithm for the purpose of decomposition. This decomposition is then used for CV selection to reduce computational time. Secondly, metaheuristic algorithms are coupled into one framework referred to as the multiagent optimization programming which is utilized to solve CV selection optimization. Lastly, real time optimization is examined for further optimality of energy plants.

Further studies should be carried out with nonlinear plants of sizes considerably larger than considered here for the purpose of control structure design. This would necessitate the utility of the estimation of structural connectivity for the purpose of reorganization of the process into different islands/sections while seeking for the optimal CV selection with the new organization.

While the DCM was developed for continuous time systems with full model integration, analysis could be extended to discrete time systems with comparison to the findings here for continuous time systems as most processes are inherently modelled as discrete systems. It may also be crucial to identify other filtering techniques that could be exploited such as particle filtering and/or unscented Kalman filters for the purpose of powerful system identification techniques. This may compromise computational efficiency. Additionally, the DCM can be utilized for obtaining pertinent information for controller design such as Gramian and Relative gain arrays.

Studies could be carried out for further algorithm development towards performance and computational time improvements. In particular, the agent-based nature of the proposed algorithm could be investigated in details for further advancements. For example, the computational time performance of the biologically-inspired methods could be improved by examining parallel computation of agent’s trajectories.

The real time optimization considers two pertinent variables for the advanced energy plant with CO₂ capture, it should be investigated if additional variables could be incorporated into real time optimization. Additionally, the methods developed in this thesis could be extended to other applications for the purpose of comparison and future research.

The biomimetic control structure design methods could be implemented further to address other cyber physical systems.
A.1 Supervisory Control Layer Design

A.1.1 Interaction Analysis

As noted earlier, three major Gramian-based measures for input-output variable interaction are the PM (Conley and Salgado, 2000), HIIA (Wittenmark and Salgado, 2002) and $\Sigma_2$ measure (Birk and Medvedev, 2003). The traditional measure for interaction, RGA, is given by Eq. (1) where $G$ is the steady-state gain and ‘.*’ denotes element-by-element matrix multiplication. In Eq. (1), the element $\lambda_{ij}$ corresponds to $y_i$ and $v_j$. Eq. (2) is the formal definition of what the elements of the RGA represent. Each of these elements shows how the gain of input $j$ on output $i$ changes when all remaining loops are closed. This provides information on loop-loop interactions as the further away an element is from 1, the higher the degree of loop-loop interactions.

$$\Pi(G) = G(0) \cdot (G(0)^{-1})^T$$  \hspace{1cm} (1)

$$\Pi(G) = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \end{bmatrix}$$

$$\lambda_{ij} = \frac{\left(\frac{\partial y_i}{\partial v_j}\right)_{\text{all loops open}}}{\left(\frac{\partial y_i}{\partial v_j}\right)_{\text{loop } i \text{ open}}}$$  \hspace{1cm} (2)

Gramian-based interaction measures are a relatively new and potentially powerful tool in the analysis of multiple-input, multiple-output control structures. The main features of these interaction measures are outlined. Readers interested in a more thorough examination are directed to (Halvarsson, 2008, Van De Wal and Jager, 2001).

The Gramian-based interaction measures all rely upon the controllability and observability Gramians. Consider the following continuous time-invariant state-space model:

$$\dot{x}(t) = Ax(t) + Bv(t)$$

$$y(t) = Cx(t)$$  \hspace{1cm} (3)
where $x(t)$ is the state vector, $v(t)$ is the input vector, and $y(t)$ is the output vector. The controllability and observability Gramians for this system are defined by Eqs. (14) and (15).

$$P = \int_0^\infty e^{A\tau}BB^Te^{A^T\tau}d\tau$$ (4)

$$Q = \int_0^\infty e^{A\tau}CC^Te^{A^T\tau}d\tau$$ (5)

The three Gramian-based interaction measures discussed in this paper all are based upon the Hankel matrix, defined as the product of the observability and controllability Gramians. An important property of the Hankel matrix is that it is independent of the state-space realization and, therefore, so is any interaction measure derived from it. Eqs. (6), (7) and (8) define the PM, HIIA, and the $\Sigma_2$ interaction matrices, respectively. The matrix norms are defined in Eqs. (9) and (10).

$$[\Phi]_{ij} = \frac{\text{tr}(P_iQ_j)}{\text{tr}(PQ)}$$ (6)

$$[\Sigma_H]_{ij} = \frac{\|P_iQ_j\|_H}{\sum_{k,l}||P_kQ_l||_2}$$ (7)

$$[\Sigma_2]_{ij} = \frac{\|P_iQ_j\|_2}{\sum_{k,l}||P_kQ_l||_2}$$ (8)

$$\|G\|_H = \sqrt{\lambda_{\text{max}}(G)}$$ (9)

$$\|G(s)\|_2 \equiv \sqrt{\sum_{i,j} \int_0^\infty |g_{ij}(\tau)|^2 d\tau}$$ (10)

An important characteristic of the Gramian-based interaction measures is that they are scaling dependent. Therefore, before these measures can be used for the design of a control structure, a systematic means of scaling must be defined. Several scaling methods have been proposed in the open literature (Salgado and Conley, 2004, Shaker and Stoustrup, 2013). In this paper, the Gramian-based interaction measures are scaled in such a way that the sum of any row (column) is equal to the sum of any other row (column). Scaling in this way ensures that all output variables are considered of equal importance, i.e., one output variable is not considered more significant than others. Additionally, it is assumed that the relative ‘power’ of all input variables is the same,
specifically, all input variables have the same relative gain. Scaling in this manner gives the Gramian matrices some of the similar properties to that of the RGA.

As mentioned earlier, the expected computation time required for the calculation of control actions if an MPC were used is proposed here as a quantitative measure of controller complexity. For simple PID controls, it is assumed that the computation is completed instantaneously and therefore has zero controller complexity. For MPC control, it is assumed the computational time and controller complexity is defined as Eq. (11). This measure is based upon the time complexity of the evaluation of an $n$ dimensional optimization problem (Karmarkar, 1984). Using this measure for controller complexity, the optimization problem shown as Eq. (12) is formulated where $J_{\text{control}}$ is calculated from one of the Gramian interaction measures. The solution of this optimization problem will yield a set of Pareto optimal control structures which balance the tradeoffs of control performance with control complexity.

$$O(n^2 \ln(n))$$  \hspace{1cm} (11)

$$\min_{v,y} \left( J_{\text{control}}(v, y) \cdot (v + y)^2 \ln(v + y) \right)$$  \hspace{1cm} (12)

A.1.2 Optimal Tuning

In addition to control structure design, tuning of MPC controllers has been a subject of ongoing research for several decades now. These tuning methods fall into one of two general categories: online and offline tuning. For a review of many of tuning methods proposed, readers are directed to (Garriga and Soroush, 2010). The method proposed in this work is an offline tuning method where the tuning parameters of the MPC are manipulated to optimize the sum of a time domain control performance metric, the integral squared error (ISE), scaled based upon the individual CV's impact on the economic performance of the process. This is a promising method, as the framework allows for the introduction of constraints on the process response and the incorporation of economic insights of the process that were attained during the course of the plant-wide control system design procedure into the tuning method.

$$\sum_{p=1}^{N_p} (\hat{y}_{k+p|k} - r)^T \Psi (\hat{y}_{k+p|k} - r) + \sum_{j=1}^{N_{\text{c}}-1} \Delta \hat{p}_{k+j|k}^T \Phi \Delta \hat{p}_{k+j|k} \text{ s.t.} \quad \hat{y}_{\text{min}} \leq \hat{y}_k \leq \hat{y}_{\text{max}}$$  \hspace{1cm} (13)
Consider a general MPC formulation, as defined in Eq. (13). Here, \( \hat{y}_{k+p|k} \) represents the vector of the plants CVs at the \((k + p)\)th time interval. Similarly the vector \( \hat{v}_{k+j|k} \) denotes the future values for the manipulated variables at the \((k + j)\)th time interval which are to be optimally decided in the face of constraints to drive the CVs \( \hat{y} \) to the reference set point \( r \) passed down from the scheduler (denoted by \( u \) at the scheduler level). The scalars \( k, p \) and \( j \) represent time, indexes for the prediction and control horizons respectively. \( N_p \) and \( N_c \) represent the prediction horizon and the control horizon, respectively. \( \Psi \) and \( \Phi \) are weighting matrices. The effects of disturbance \( \theta_k \) at any time \( k \) is incorporated into the discrete state space model.

The ‘tuning parameters’ for this MPC are the sampling interval, \( T_s \), the prediction horizon, \( N_p \), the control horizon, \( N_c \), and the weighting matrices \( \Psi \) and \( \Phi \). In this work, the prediction horizon is set following the heuristics of (Banerjee and Shah, 1992) to a value of 95% of the settling time to steady state and the control horizon is set following the heuristics of (Georgiou et al., 1988) to a value of 60% of the settling time to steady state.

For the determination of the optimal output and movement suppression weights, an optimization problem is formulated. Here, \( n_y \) is the number of CVs, ISE is the integral squared error of the primary controlled variable \( i \), and \( \Theta_i \) is the scaling factor based upon the economics of the process. The objective function to be optimized is the summed, scaled ISE values of the CVs, defined as in Eq. (14). The scaling factors, \( \Theta_i \) are based upon the impact of individual CVs on the economics of the process. These are the same scaling factors as used in our previous work (Jones et al, 2014) for the selection of secondary CVs.

\[
\min_{\Psi, \Phi} \sum_{i=1}^{n_y} \Theta_i ISE_{y_i} \quad (14)
\]

\[
\text{Subject to} \quad \Gamma(r, y, v, \ell) \leq 0 \quad (15)
\]

In addition to the minimization of the summed, scaled ISE, one can include constraints on the process responses given set point changes or measured/unmeasured disturbances. For example, one may wish to specify that a controlled variable have no more than a 3% overshoot in response...
to a step change to its set point. Including such constraints within the optimization allows for important process characteristics to be addressed during tuning the initial tuning of the MPC. These inequality constraints can take many user defined forms and represented as Eq. (15).

Results

First, the optimal structure of the supervisory control layer needs to be selected as outlined in A.1. To begin, the state space model of the AGR unit is required. This is obtained from the Aspen Plus Dynamics model of the AGR unit. From this state space model, the controllability and observability Gramians are calculated for each of the individual subsystems, i.e., each of the pairings of input to output. From these calculations, the three unscaled Gramian interaction matrices are obtained. Next, each of these Gramian interaction matrices are scaled, according to the methodology discussed in Section A.1. The three Gramian interaction matrices, namely HIIA, PM and $\Sigma_2$ interactions measures, are used to determine the optimal pairings of the structure. These Gramian interaction measures may lead to the same or different control structures. The optimization problem shown in Eq. (9) is solved for all possible control structures that involve either decentralized or centralized, or any combination thereof where $J_{control}$ is calculated from the Gramian interaction measure used. To determine the actual structure to be used, the numerical derivative of the control performance criteria with respect to the controller complexity is calculated for the PM, HIIA and $\Sigma_2$ interactions measures listed in Table A.1. From this table, it is observed that using a combination of ‘one 4 by 4 centralized’ controller and ‘one 2 by 2 centralized’ controller is optimal.

Table A.1 Numerical Derivative of Control Performance with respect to Controller Complexity [listed in increasing controller complexity]

<table>
<thead>
<tr>
<th>Disturbance</th>
<th>$\Sigma_2$</th>
<th>PM</th>
<th>HIIA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decentralized</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>One 2x2 Centralized</td>
<td>0.039638</td>
<td>0.036428</td>
<td>0.040864</td>
</tr>
<tr>
<td>Two 2x2 Centralized</td>
<td>0.034841</td>
<td>0.033038</td>
<td>0.030405</td>
</tr>
<tr>
<td>Three 2x2 Centralized</td>
<td>0.021244</td>
<td>0.018322</td>
<td>0.012263</td>
</tr>
<tr>
<td>One 2x2 Centralized One 3x3 Centralized</td>
<td>0.012298</td>
<td>0.015821</td>
<td>0.014232</td>
</tr>
<tr>
<td>One 4x4 Centralized</td>
<td>0.010963</td>
<td>0.010176</td>
<td>0.009009</td>
</tr>
<tr>
<td>Two 3x3 Centralized</td>
<td>0.01767</td>
<td>0.011027</td>
<td>0.006856</td>
</tr>
<tr>
<td>One 4x4 Centralized One 2x2 Centralized</td>
<td>0.002028</td>
<td>0.00617</td>
<td>0.007597</td>
</tr>
<tr>
<td>One 6x6 Centralized</td>
<td>0.009358</td>
<td>0.007953</td>
<td>0.006746</td>
</tr>
</tbody>
</table>
With the structure of the supervisory control determined, the design of the supervisory control system is undertaken. The 4x4 centralized controller comprises the following controlled variables, namely CO₂ capture rate, vapor composition in the CO₂ absorber, H₂S purity to the Claus unit, and solvent composition in the H₂S absorber, and the following manipulated variables, namely the LP flash pressure, semi-lean solvent flowrate, lean solvent flowrate, and H₂S concentrator pressure. The 2x2 centralized controller controls H₂S capture and water content of the solvent using the stripper bottom temperature and steam flow to the stripper. For the purposes of this work, the forms used for these centralized controls are linear model predictive controls (LMPC).

Models were identified by applying a pseudorandom binary sequence (PRBS) input signal to the nonlinear process model in Aspen Plus Dynamics. Using the MATLAB system identification toolbox, the output data and the PRBS input data were used to identify linear transfer functions. Using these identified models, the LMPCs for the process are designed. For the tuning of the LMPCs, the economic information obtained during controlled variable selection is introduced as described in (Jones et al, 2014). The objective of the optimization is shown in Eq. (11). Table A.2 shows comparison of the objective function values of the initial, non-optimized tuning used for the LMPCs and that of the PID controllers. Table A.2 shows that superior performance, as compared to PID control, is attained from the LMPCs using these tuning parameters.

**Table A.2** Comparison of Initial ISEs of the LMPC to PID for Three Disturbances

<table>
<thead>
<tr>
<th>Disturbance</th>
<th>Integral Square error (ISE)</th>
<th>Percent improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PID</td>
<td>LMPC</td>
</tr>
<tr>
<td>-20% Step in syngas flow</td>
<td>701.4</td>
<td>222.7</td>
</tr>
<tr>
<td>+2% Step in CO₂ Capture</td>
<td>116.9</td>
<td>43.0</td>
</tr>
<tr>
<td>-2% Step in CO₂ Capture</td>
<td>103.7</td>
<td>49.3</td>
</tr>
</tbody>
</table>

For the first row, comparison of the performance of the PID controller and MPC for CO₂ capture, H₂S purity to Claus unit, CO₂ vapor fraction in CO₂ absorber and the scaled ISEs is shown in Figs. A.1 and A.2, respectively, for a -20% step change in syngas flow.
Figure A.1 CO₂ Capture Fraction (a) and H₂S Purity to Claus unit (b) after 20% Step Decrease in the Syngas Flowrate

Figure A.2 CO₂ vapor fraction (a) and ISEs (b) due to 20% Step Decrease in the Syngas Flowrate
A.2 Dynamic Causal Model

Table A.3 Latent connectivity for the acid gas removal unit

<table>
<thead>
<tr>
<th>Variables</th>
<th>CO$_2$ absorber (T1)</th>
<th>H$_2$ recovery KO drum (D1)</th>
<th>H$_2$ recovery drum (D2)</th>
<th>MP flash (D3)</th>
<th>LP flash (D4)</th>
<th>H$_2$S absorber (T2)</th>
<th>H$_2$S concentrator (T4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO$_2$ absorber (T1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ recovery KO drum (D1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ recovery drum (D2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MP flash (D3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LP flash (D4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S absorber (T2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S concentrator (T4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$ liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S vapor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H$_2$S liquid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Normalization

Consider the equation

\[
\dot{X}(t) = \left( \bar{A} + \sum_j u_j \bar{B}^j + \text{diag}(X) \bar{H} \right) X
\]  

(16)

Let the variables be scaled such that

\[
\dot{\bar{X}}(t) = \left( \bar{A} + \sum_j \tilde{u}_j \bar{B}^j + \text{diag}(X) \bar{H} \right) \bar{X}
\]

(17)

Where

\[\bar{X}_i = X_i / \max(X_i)\]

\[\tilde{u}_p = u_p / \max(u_p)\]

It follows that

\[\bar{A}_{ij} = \frac{\bar{A}_{ij} \max(X_i)}{\max(X_j)}\]

\[\bar{B}_{ij}^p = \begin{cases} \frac{\bar{B}_{ij}^p \max(X_i)}{\max(u_p) \max(X_j)} & \forall j > 1 \\ \frac{\bar{B}_{ij}^1 \max(X_i)}{\max(u_p)} , j = 1 \end{cases}\]

\[\bar{H}_{ij} = \frac{\bar{H}_{ij}}{\max(X_j)}\]
References


AL-SINBOL, G. 2013. Analysis of gps abnormal conditions within fault tolerant control laws, West Virginia University.


GARCIA, D. 2010. Robust smoothing of gridded data in one and higher dimensions with missing values. Computational statistics & data analysis, 54, 1167-1178.


