ANALYSIS AND DESIGN USING MODULARITY CONCEPTS IN CHEMICAL PROCESSES AND SUPPLY CHAINS

by

YUE SHAO

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Chemical Engineering)

at the
UNIVERSITY OF WISCONSIN-MADISON
2022

Date of final oral examination: May 17, 2022

The dissertation is approved by the following members of the Final Oral Committee:

Victor M. Zavala, Baldovin-DaPra Professor, Chemical and Biological Engineering
Carla Michini, Assistant Professor, Industrial and Systems Engineering
Ross E. Swaney, Associate Professor, Chemical and Biological Engineering
Styliana Avraamidou, Assistant Professor, Chemical and Biological Engineering

To my family.

ACKNOWLEDGMENTS

I can't believe it has been almost 5 years since the first day I got here. I was even reluctant to go outside when I came to Madison the first time during the school visit from Houston, my cozy and warm home at that time. But here I am now, a proud badger that survived one of the harshest winter here and that enjoyed every minute at this beautiful town in the north during this 5 years of my Ph.D. program. It has been an unforgettable journey in my life thanks to all the wonderful people I met.

First of all, I want to express my sincere and greatest gratitude to my advisor, Prof. Victor Zavala for "babysitting", mentoring and trusting me through my study in the graduate school. Switching from lab work (during my undergraduate research) to computing and optimization wasn't easy for me initially, and his feedback could be critical sometimes in the beginning. However, he never gave up on me and these words made me realize that I needed to work harder to bridge the gap. He babysat me for my first publication when I had no idea how to compile a scientific paper, guided me for the direction of my research when I felt lost along my way, and most importantly supported and encouraged me whenever I needed help for my research or in my life. Knowing that I would like to go to the industry after my graduate study and how hard it could be for international student, he referred me for an internship opportunity during the summer of my third year, and supported me on learning new knowledge that are not directly related to my research and doing an internship based on it. I would like to not only express my great appreciation for his technical support in the field of mathematical modeling, optimization, and data science, but also show my respect for him being an extraordinary leader, an adviser with responsibilities, and an excellent communicator. The journey was rugged and uncertain at fist, but now, thanks to him, I am finishing my Ph.D. dissertation.

I would like to thank Prof. Ross Swaney, Prof. Styliana Avraamidou, Prof. Carla Michini for serving on my final defense comittee, giving me the opportunity to present my work, and sharing valuable feedback. I would also express my appreciation to Prof. Dan Klingenberg, Prof. Opheila Venturelli and Prof. James Dumesic for being committee members in my preliminary exam, providing early suggestions and direction on my research. Special thanks to Prof. Laurent Lessard (now a professor in Northwestern University) who brought me into the world of optimization and got me started on Julia programming language during the Introduction to Optimization course, and Prof. James Luedtke in the department of Industrial and Systems Engineering who had taken time to support me for job and academic applications. I would also like to thank the depart-

ment coordinator Kate Fanis for providing administrative help along my way, and the technical support I received from Russ Poyner, our information process consultant in the department. They all made my Ph.D. study a lot smoother.

I would also like to thank all members in the scalable systems laboratory, or *zavalab*. I had great memories of getting group lunch on Fridays, chatting in the office, and attending conferences with you all. I want to thank Apoorva Sampat, Ranjeet Kumar, Yankai Cao, Jordan Jalving, Yicheng Hu, and Sungho Shin for all your help when I first joined the group. I also want to thank Alexander Smith, Joshua Pulsipher and all other members for being amazing lab-mates and friends, and creating an easy and fun group environment.

I also feel fortunate to have the opportunity to work with Lav Thyagarajan at John Deere Electronic Solutions despite the severe Covid conditions that showed me how it was like to work in the industry, with Alan Strancar and Haibin Chen at Proctor & Gamble for guiding through data science and machine learning projects, and finally with Braulio Brunaud at Johnson & Johnson for working as a process and systems engineer. All these industry working experience were essential for me to explore what I really want to do after my graduate school and eventually helped me get my offer at Dow in Houston.

I also want to mention my incredible friends in Madison who made the journey more enjoyable. I would like to thank Zhijian Tao for helping me adopting a healthier life habit and lose 50lbs during the first couple of years here, and Zhe Yang, and Wenyuan Zhu for hitting the gym with me and bearing with me this whole time. My roommate, Zheyu Su, has always been considerate and easygoing, and I enjoyed my last three years living with him. I would also like to thank Xinyi Liu, Yijiang Yu, Na Li, Zhaoning Yu, and Zhuoyan Xu for providing constant support, and Xuan Wang and Wei Zhao for providing professional advice during my job search. And I want to thank Yuan Fang, Haotian Jiang, Yajing Ye, Yu Chen, Siying Qu and all other friends for your accompany in Madison.

I would have never had the opportunity to purse my graduate studies in Madison without the excellent education I received during my undergraduate study at Rice University. I was exposed to fours years of well-rounded chemical engineering curriculum. Prof. George Hirasaki, Prof. Francisco Vargas, and Prof. Sibani Biswal inspired me to pursue my graduate studies in chemical engineering. I also made life-long friends at Rice University who continue to support me even after almost a decade.

Last but probably most importantly, it would have been impossible to achieve any of this without my family. They have supported me to pursue my dreams abroad after I turned 18, and have always provided me with the best of everything I ever needed along the way. As the only child in my family, I want to express my deepest gratitude and love to my Mom, Rong Gao, and my Dad, Yuqiang Shao, for giving all of their unconditional love and support to me. I would also want to thank my cousin Yuan Shao for the life advises provided as a big brother. I could not have achieved this without you all.

Yue Shao
Madison, WI
May 2022

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Modularity concepts have been recently explored in the context of industrial production (manufacturing) systems such as chemical processes, energy systems, and infrastructures. Standardization and size reduction brought by modular technologies enable mass off-site fabrication, fast transportation, and deployment of equipment, which could ultimately lead to technology cost reductions. Modular systems, contrasting with large and centralized systems that involve lengthy on-site construction phases and difficult transportation (and are thus rarely relocated), also enable sequential investment strategies that provide flexibility to mitigate market and regulatory risks. Therefore, a thorough study of modularity concepts using computational frameworks and optimization tools is required.

We study modularity in the context of manufacturing systems and address the following three questions: what is modularity (concepts), why should we consider modular designs (benefits), and how can we incorporate modular concepts in actual designs (usage). Specifically, we first propose a modularity measure that captures unique characteristics in the context of manufacturing systems. Then, we study the spatial and temporal flexibility brought by modular construction (decentralization in electricity market and capacity expansion planning). Finally, we propose a system-level design concept: spatial superstructure, and based on this new concept, we derive a formulation that provides optimal spatial modular design of a process or supply chain.

By providing a complete story on modularity in manufacturing and supply chain, we seek to encourage incorporating modular concepts in future process and supply chain designs. In this work, we provide necessary definitions and computational frameworks that aid the modular process design and solve potential design problems such as scalability.

1

INTRODUCTION

1.1 Traditional Centralized Chemical Plants

In the past century, two rules have defined and guided the process system design in modern chemical engineering: the economy of scale and unit operations (Baldea et al., 2017). First, economies of scale are the cost advantages that businesses obtain as a result of their size of operations and can be described as the decrease in cost per unit of output due to an increase in scale (Silvestre, 1987). In chemical engineering, constructions of chemical plants can be more capital efficient by simply making them larger, reducing the operational cost and products prices and thus improving the profits (Berthouex, 1972). A prominent example of it in chemical engineering is the cost of pipes (openstax). The cost of materials used to manufacture a pipe is proportional to the pipe's circumference and length. However, the cross-sectional area of a pipe determines the volume of chemicals that may flow through it. Thus, a pipe which uses twice as much material to make can carry four times the volume of chemicals because the cross-section area of the pipe rises by a factor of four. Of course, the actual economies of scale for chemical plant are far more complicated than what this example suggests, but process engineers and designers have long used the so-called "six-tenths rule"(or "two-thirds rule"), a rule of

thumb which indicates that increasing the amount of product produced in a chemical plant by a given percentage will only result in a six-tenth (or two-thirds) rise in overall cost (Kerridge, 1982). Second, in the early 20th century, Warren K. Lewis, one of the pioneers in chemical engineering, together with his several colleagues explained in their book, The Principles of Chemical Engineering, that the variety of chemical industries have processes that follow the same physical laws, and they concluded these similar processes into unit operations to facilitate calculations in an era ruled by the slide rule rather than by the digital computer (Walker et al., 1923; Cabe et al., 2018). While slide rules have become obsolete nowadays, unit operations remained as a central piece in the design, simulation and optimization of chemical processes.

In the past two decades, these two principals - economies of scale and unit operations - have undergone serious challenges as markets become more and more volatile and fast-paced, and resources become more sparsely scattered (Gwehenberger and Narodoslawsky, 2008; Stevens, 2003; Charpentier, 2002). Based on the electricity price data for more than 5000 locations in California throughout the year 2015 collected from California Independent System Operator (CAISO), the average electricity price for real time market (RTM) is 32.71 USD/MWh while the temporal volatility is 62.42 USD/MWh (Shao and Zavala, 2019). And in one particular location, the price shifts from 50 USD/MWh to 600 USD/MWh and then back to 35 USD/MWh in just 20 minutes. If a conventional ammonia manufacturing plant which consumes the power at 50 MW was built at this location (Egenhofer et al., 2017), the risk caused by electricity price fluctuation is enormous. What is more, building a chemical plant of this size usually takes 5 to 7 years of on-site construction (Rahman et al., 2014). During this period of time, if the initial demand forecast becomes overly optimistic (especially when some unforeseen events such as COVID-19 happen) or a new type of technology or product appears to replace the old one, investors are likely to lose billions of dollars due to this long project time-line (Roy and Eng, 2017). Eventually, transporting raw materials to a centralized plant for processing might not be possible for certain industries. For instance, many oil gathering companies have been

flaring and venting off natural gas in the Permian Basin of west Texas (Davis, 2016). Flaring is a well-head technique used to dispose of excess natural gas that is produced along with oil (Emam, 2015). According to the U.S. Energy Information Administration, a trillion cubic feet of natural gas in the Permian Basin has been flared since 2013 and it is responsible for 25 percent of today's warming from human activities. It occurs because the pipeline infrastructure could not carry the amount of natural gas, and burning it is the most cost efficient way (Udok and Akpan, 2017; Soltanieh et al., 2016). Capturing or processing the flare gas by building traditional centralized plant is not viable as the plant will be abandoned when the oil well dries out. To solve the above mentioned challenges that are specific to today's markets and environmental concerns, new design concept such as modular design (or modularization) has become prevalent (Roy and Eng, 2017).

1.2 Modular Design Concepts/Modularization

Modularization is a manufacturing trend that is being adopted in different industrial sectors such as power generation, data centers, chemical processes, and supply chains (Frivaldsky et al., 2018; Berthélemy and Rangel, 2015; Dong et al., 2009; Chakraborty et al., 2009; R., 1999). For instance, decentralized power generation and storage systems are becoming increasingly attractive as climate changes and adoption of renewable power disrupt markets and space-time demand patterns (Heuberger et al., 2017; Liu et al., 2018; Shao and Zavala, 2019). Modular manufacturing systems are typically built from small-scale and standardized technologies (equipment units) that perform self-contained tasks and that are coupled together using sparse interfaces. Small dimensions and sparse interfaces facilitate system assembly/dis-assembly and reconfiguration (e.g., migration of technologies to a different location and expansion of capacity). This logistical flexibility helps systems adapt to fast-changing markets and other externalities (e.g., climate, resource availability, policy) (Jaikumar, 1986; Rajagopalan, 1993) and enables the recovery of resources that are highly distributed and potentially short-lived (Allman and Zhang,

2020; Chen and Grossmann, 2019; Davis, 2016). In the context of chemical engineering, a modular plant can be comprised of a single module that incorporate many unit operations or of multiple modules that are shop-fabricated separately, transported to the manufacturing site, and eventually connected to form a large process system (Bieringer et al., 2013; Kockmann, 2016). The process equipment, valves, piping components, and electrical wiring are installed within a structural steel framework known as a module and each module is a self-contained process unit that is typically built off-site (Seifert et al., 2012). The project start-up time for a modular plant is minimized since systems can be fully assembled and tested before they ship, reducing the amount of on-site construction time (Kockmann et al., 2017). Lower labor and operational costs are achieved due to a shorter project time-line, efficient use of material, and a smaller field crew (Hady et al., 2009). Also, off-site module construction does not interrupt or shut-down pre-existing operations. All these advantages have made modular design a preferred choice that adapts today's market requirements and in fact, many pioneering companies have announced their modular solutions for chemical engineering industries.

Figure 1.1: "Flarecatcher" by Pioneer Energy (Pionner Energy)

Pfaudler Group, a leading provider of technologies, systems and services for the chemical and pharmaceutical industries, designs complete modular process systems with

all ancillary equipment, piping, and instruments included. Pfaudler is able to design and construct a large Wiped Film Evaporator (WFE) system as two large modules within two months and then complete the on-site construction of them within 24 hours. Pfaudler leads the way in providing complete process systems as pre-assembled modules, which provide lower project costs, shorter construction schedules, and improved quality (Pfaudler Group). Another example that utilizes modular concepts is the "Flarecatchers" constructed by Pioneer Energy. At well sites or central processing facilities, "Flarecatchers" extract Natural Gas Liquids (NGLs) from rich, associated and non-associated gas while conditioning the remaining gas. The remaining gas is then conditioned to pipeline quality, making it ideal for on-site power generation, conversion to compressed or liquefied natural gas, or injection into a lean gas pipeline. As shown in Figure 1.1, their systems are skid-mounted, modular, autonomous units that are remotely monitored and controlled. This enables flexibility in equipment deployment and superior up-time, while minimizing required capital and operating expenditures (Pionner Energy). The merits of modular design concept are not only suitable in process system design, but also applicable in areas such as supply chain design (Becker et al., 2021; Doran and Giannakis, 2011; Palys et al., 2018), and many industrial companies are switching from traditional centralized plants to this more efficient, flexible, and environmental friendly way of design and construction (Epic Systems).

1.3 Research Questions on Modular Designs

The concept of modular design is not difficult to understand, and the benefits of it seem to be obvious. However, the adaptation of the concept is slow, and many companies still hesitate on the lack of quantitative analysis and mathematical backup of the idea. Therefore, a thorough analysis and study using mathematical tools on modular process and system design is needed. First, the definition of modularity or modularization is ambiguous in the context of manufacturing. In graph theory, for a given organization,

the amount of internal module coupling relative to coupling between modules is referred to as the degree of modularity (Newman, 2006; Brandes et al., 2007). While the notion of modularity is pervasive in all areas of science and engineering (Langlois, 2002; Coltheart, 1999; Baldwin et al., 2000), there are no well-established mathematical (quantifiable) measures of modularity in manufacturing. The availability of such measures is key to enable more systematic analysis and design of modular systems. Second, quantitative analysis on the benefits of modular design is missing. We would like to provide a quantitative view on the flexibility, cost saving, and risk minimizing effects brought by modular design comparing to the traditional centralized design. This direct comparison would provide tangible proofs and build mathematical foundations for the benefits of modular design concept. As a final step, we would like to extend the modular concept to design modular supply chains and propose a general computational framework that provides tools to guide the synthesis of modular processes and supply chains. By providing a complete story on modularity in manufacturing and supply chain, we seek to encourage incorporating modular concepts in future process and supply chain design, and provide necessary definitions and computational frameworks that aid the modular process design and solve potential design problems such as scalability.

By identifying these issues, this work tries to provide insights to the following research questions about quantifying modular design concept and providing modular design schemes:

1. What - What is modularity? How do we define a quantifiable measure for modularity in manufacturing context?

2. Why - Why should we consider modular design? Can we use mathematical tools to demonstrate the benefits of modular designs comparing to traditional ones?

3. How - How do we design a modular process or supply chain? Can we provide a general framework to aid the synthesis of modular processes and supply chains?

1.4 Outline

This dissertation is structured as follows:

Chapter 2 – Modularity Measures: Concepts, Computation, and Applications to Manufacturing Systems. This chapter addresses the "What" question. It introduces a measure to quantify the modularity of industrial production (manufacturing) systems and proposes an optimization formulation to compute it. It provides a discussion on advantages and disadvantages of alternative modularity measures used in different scientific and engineering communities and the uniqueness of our proposed measure in manufacturing context.

Chapter 3 – Benefits of Modular Design - Spatial Flexibility This chapter addresses the first part of the "Why" question. In this chapter, we study the economic incentives for deploying modular technologies created by space-time dynamics of day-ahead and real-time electricity markets. We first formulate the electricity market clearing model and explore the effects of modularization on social welfare and electricity prices. We then develop an optimal technology placement formulation that seeks to identify optimal strategies to maximize expected profit and minimize risk in the electricity markets.

Chapter 4 – Benefits of Modular Design - Temporal Flexibility This chapter addresses the second part of the "Why" question. This chapter studies logistical investment flexibility provided by modular processing technologies for mitigating risk under the setting of the capacity expansion problem. This capacity expansion problem is a stochastic, multi-stage, and multi-objective optimization problem, and we study the trade-offs between expected profit and risk for capacity expansion plans with different sizes of technologies available. Case studies of different complexity are presented to illustrate the developments.

Chapter 5 – A Spatial Superstructure Approach to the Optimal Design of Modular Systems. This chapter addresses the "How" question. It first introduces a design concept, Spatial Superstructure, that encodes spatial (geographical) context of all system components. We then extends the design concept to include the modularity measure proposed in previous section, which enables the simultaneous modular design of processes, facilities, and of supply chains. Finally, we propose an optimization framework based on this new design concept to facilitate the design of modular manufacturing systems.

Chapter 6 – Conclusions and Future Directions. This chapter provides a summary of the major contributions of this dissertation. Specific areas for improvement for future research are also identified.

2

MODULARITY MEASURES: CONCEPTS, COMPUTATION, AND APPLICATIONS TO MANUFACTURING SYSTEMS

This chapter is published in Shao and Zavala (2020).

2.1 Introduction

Modularization is an organization strategy that is used in living, socio-economic, and industrial systems to facilitate learning and evolution and to cope with complexity (Simon, 1962; Langlois, 2002). For instance, biological networks and the human body exhibit high modularity (Ravasz et al., 2002; Meunier et al., 2010; Newman, 2006; Meunier et al., 2009). This organization structure facilitates specialization of components (e.g., organs and metabolic cycles) and enables management of large numbers of functions. In a modular organization, fundamental components and associated functions are grouped into clusters (modules). Modules have the distinctive feature that coupling between internal components (intra-module) is significantly stronger than coupling across modules (inter-module). Scientists have long argued that modular organization provides flexibility and facilitates evolution because modules can adapt, mature, or disappear without significantly disrupting the entire system. This arrangement also facilitates the management of

complexity because tasks and information are refined progressively. Herbert Simon, one of the pioneers of computer and cognitive science, argued that it is rather natural that human-made organizations (e.g., government institutions and enterprises) also exhibit high degrees of modularity (Simon, 1962). This is because the human brain processes information and makes decisions in a modular manner (Meunier et al., 2009). Modularity provides an indication of the flexibility and maturity of an organization and of the range of functions that it can perform (Langlois, 2002). Modularity has also been found to facilitate the control of large networks (Constantino et al., 2019).

Modularity concepts have also been recently explored in the context of industrial production (manufacturing) systems such as chemical processes, energy systems, and infrastructures. Industrial production systems can be built from small-scale and standardized equipment modules that perform well-defined tasks and that are coupled together using well-defined and sparse interfaces (Seifert et al., 2012; Bramsiepe et al., 2012; Baldea et al., 2017). Standardization and size reduction enables mass off-site fabrication and fast transportation and deployment of equipment, which accelerates experimentation and learning and ultimately leads to technology cost reductions (Lier and Grünewald, 2011; Lier et al., 2016; Rogers and Bottaci, 1997; Tatum, 1987; Roy and Eng, 2017; Hesler, 1990). A celebrated example of this principle is Henry Ford's assembly line (Langlois, 2002). Modular systems contrast with large and customized systems, which involve lengthy on-site construction phases and difficult transportation (and are thus rarely relocated); this systems also provide limited experimentation/testing opportunities (Wells, 1979). Modular systems also enable sequential investment strategies, which provide flexibility to mitigate market and regulatory risk (Hagspiel et al., 2016). Small modular systems can also facilitate the processing of geographically dispersed resources that are deemed too expensive to collect and centralize.

Modularization can accelerate investment in technologies such as small nuclear reactors, distributed generators, power electronics, chemical processes, and battery storage systems (Berthélemy and Rangel, 2015; Chakraborty et al., 2009; R., 1999; Rothwell, 2006;

Dong et al., 2009). Specifically, large industrial facilities (reaching investments of billions of U.S. dollars) might involve slow deployments and risks that few investors are willing to tolerate. On the other hand, modularization provides flexibility in investment size and enables faster deployments that ultimately result in reductions in time-to-market and facilitates financing. Moreover, expansion of production capacity in modular systems can proceed sequentially, which provides a mechanism to hedge against the market and regulatory risk. We can interpret the ability to accelerate and stage investment (and thus a hedge against risk) as a form of built-in logistical temporal flexibility. Modularity can also provide logistical spatial flexibility in the sense that small modules can be easily transported and relocated. This can enable the recovery of resources that are highly distributed and potentially short-lived. As a result, it has been argued that modularization can enable more sustainable systems and circular economies (Seliger and Zettl, 2008). For instance, modular systems can be used to harness natural gas resources that remain stranded at oil production facilities due to limited gas pipeline infrastructure (Davis, 2016). Modular technologies can also be used to recover bio-gas from organic waste generated at animal farms, landfills, and waste-water treatment facilities. It has also been recently observed that modular systems can be strategically placed to exploit space-time electricity price dynamics and with this mitigate risk (Shao and Zavala, 2019). In this context, module transportability is important from a relocation perspective. For instance, unlike large central systems, modular systems might not be permanently placed at a single location but might be disassembled, relocated, and re-assembled at different locations throughout their lifetime based on changes in resource availability, policy, weather, and infrastructure. For instance, a change in government regulations might render a given facility location undesirable or obsolete. Small modules can also be transported back to shops to perform maintenance and can be quickly replaced. As can be seen, space-time logistical flexibility provided by modularity can allow organizations to diversify, mitigate risk, and have a higher likelihood of surviving strong fluctuations of markets, government regulations, and other externalities. On the downside, flexibility provided by small

modular systems often comes at the expense of higher investments and reduced operational efficiency when comparing to large systems. Specifically, economies of scale benefit large systems due to the favorable scaling of throughput with equipment size. Industrial systems will thus likely evolve into a mixed state in which certain tasks are performed in small dispersed modular systems while others are performed in large centralized facilities. This has the potential of inducing a re-organization of production facilities and of entire supply chains. This is particularly the case in chemical processes and power plants (Peters et al., 1968). This reasoning also indicates that large centralized facilities are operationally efficient but logistically inefficient from an assembly and transportation stand-point.

The concept of modularity is pervasive in science and engineering but, surprisingly, there are few quantifiable measures of modularity. The availability of proper measures is key to enable more systematic analysis, design, and comparison of modular systems. In pioneering work, Newman proposed a modularity measure that quantifies the edge density of a system (represented as a graph) relative to the expected edge density of a random graph (Clauset et al., 2004; Newman, 2006). The argument behind this measure is that modular organizations that arise in natural systems are non-random. This measure is intuitive and has seen many interesting applications; for instance, this measure has been shown to provide a flexible and powerful tool for the analysis and design of control architectures (Jogwar and Daoutidis, 2017; Moharir et al., 2019; Daoutidis et al., 2018; Tang et al., 2018; Pourkargar et al., 2018, 2019) and for the decomposition of large-scale optimization problems (Moharir et al., 2017; Allman et al., 2019). A powerful generalization of Newman's measure has been proposed in (Reichardt and Bornholdt, 2006) and here it was shown that systems of high modularity are extremum points of a Hamiltonian function. Information-theoretical interpretations of modularity have also been proposed in the literature (Rosvall and Bergstrom, 2008).

Unfortunately, the modularity measure proposed by Newman (and its generalizations) do not have an intuitive interpretation from a manufacturing perspective and fail

to capture some desirable features arising in this context (e.g., module dimensions). In order to define alternative modularity measures, it is important to highlight that: modularity is not a classification but a measure (i.e., systems have different degrees of modularity). Moreover, one should note that a system with fixed physical connectivity (topology) can have different modular organizations with associated degrees of modularity and that topology dictates the number of alternative modular organizations and their associated modularity value. While these notions are clear from a conceptual point of view, there is significant ambiguity associated with the definition of modularity in manufacturing. In the metal processing industry, for instance, a module is defined as a technically and organizationally limited area of a facility that fulfills a defined task in terms of company-internal or -external salable goods and services (Wiendahl H.-P. and F., 2005). In the process industry, a module is defined as an unmodifiable element that provides a dedicated function for the process and is reusable during the planning or realization of modular plants (Hohmann et al., 2017). In other words, a module is a standardized and self-functioning unit. While these definitions are intuitive, they do not provide means to quantify modularity. Specifically, under these definitions, any equipment unit or an entire facility itself can be a module. Moreover, these definitions fail to capture aspects such as transportability and dimensions.

In this chapter, we propose measures to quantify the modularity of manufacturing systems and optimization formulations to compute them. We claim that, from a manufacturing perspective, a system is deemed modular if: i) the equipment units that compose it form clusters (modules) of dense connectivity (i.e., difficult module construction is performed off-site), ii) connectivity between modules is sparse (i.e., easy module assembly is performed on-site), iii) the number of modules is small, and iv) the module dimensions facilitate transportation. In the proposed framework, a facility has a topology that is modeled as a graph. Here, the physical equipment units represent nodes that are coupled together via edges. This representation allows us to borrow concepts and techniques from graph theory. Specifically, from a graph-theoretical perspective, the partitioning of

a graph induces a modular organization, as each partition (module) is composed of a set of nodes. For a given organization, the amount of internal module coupling relative to the coupling between modules is referred to as the *modularity* (Newman, 2006). In our approach, the proposed measure is computed for a graph by finding the partition that induces the maximum modularity (given a fixed number of modules). We show that this measure can be computed by solving a convex mixed-integer quadratic program. Additionally, we show that the mixed-integer representation allows us to impose additional features such as module dimensions and to identify multiple solutions that give the same level of modularity. We compare the proposed measure against that of Newman (widely used in other scientific communities) to highlight the advantages and disadvantages from a manufacturing perspective. This analysis reveals that the mixed-integer programming formulations proposed can also be used to compute the measure of Newman while handling constraints and that they can be used to find multiple solutions. Moreover, the proposed measure can be used within optimal design formulations and in other applications beyond manufacturing (e.g., design of control architectures and decomposition of large sets of equations).

2.2 Measures of Modularity

In our framework, we assume that connectivity, number of modules, and dimensions are key features that define the modularity of a system. Connectivity and number of modules dictate the nature and complexity of off-site and on-site assembly tasks while dimensions dictate whether modules are transportable (and thus off-site assembly is possible) and dictate economies of scale. Module connectivity is related to the problem of community detection in networks, which has been widely studied in graph theory (Fortunato, 2010). Motivated by this, it seems natural to model a manufacturing system as a graph that is composed of nodes (equipment units) and edges (that connect the units) and use graph theoretical techniques to identify modular organizations. Current techniques available

include hierarchical clustering algorithms (Trevor Hastie and Friedman, 2009; Heo et al., 2015), k-means clustering (Rattigan et al., 2007), spectral clustering (Donath and Hoffman, 1973), and techniques based on modularity maximization (Clauset et al., 2004; Newman, 2006). An excellent review on community detection techniques is provided in (Daoutidis et al., 2019). In this work, we adopt a modularity maximization approach, as this provides an intuitive approach to analyze and design modular manufacturing systems.

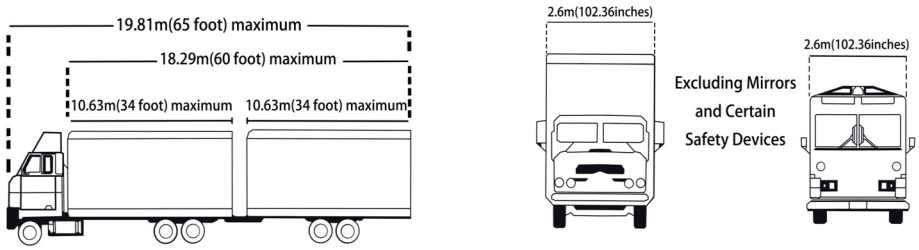


Figure 2.1: United States federal regulations on transportation dimensions (Fed, 2017).

Transportation logistics is a key factor that is often overlooked in modularity studies and that is unique to manufacturing (compared to other scientific disciplines such as neurology). In particular, a commercially-viable equipment module must be transportable using available infrastructure (e.g., railway and trucks) (Haney et al.). As such, module dimensions (length, width, and height) and weight must follow government regulations. For instance, according to the United States regulations for commercial motor vehicles, the maximum width allowed in an interstate highway is 102-130 inches (2.6-3.3 meters), the maximum height allowed is 14-16 feet (4.3-4.9 meters), the maximum length allowed is 75 feet (22.86 meters), and the maximum weight is 44,000 lbs (Haney et al.). Consequently, any system that does not satisfy these limits must be partitioned in order to enable transportation. For example, a distillation system with a diameter of more than 120 inches and height of 100 feet (around 40 trays with 24 inches spacing) must be partitioned to enable transportation and must be assembled on-site. As expected, the larger the dimensions of the system the more partitions that will be needed and the more

complex the on-site assembly.

We observe that the different features desired for modular systems might be conflicting. For instance, a small system might be the ideal modular system in that it can be completely assembled off-site (i.e., minimizing on-site assembly tasks), packed in a single module, and transported to its final destination. However, this small system might be inefficient from the perspective of economies of scale. Consequently, one might be willing to modularize only certain components of the system (thus increasing the number of modules but increasing efficiency). As another example, note that one might intentionally prefer a system with a larger number of modules in order to facilitate shop assembly of different types of modules and at different locations.

2.2.1 *Graph Theoretical Concepts*

We model a system as an undirected graph $G = (V, E)$ where V is its set of nodes (vertices) and E is its set of edges. We define the number of nodes as $n := |V|$ and the number of edges as $m := |E|$. Connectivity of nodes in G is encoded in the adjacency matrix $A \in \mathbb{R}^{n \times n}$ with entries $A_{i,j}$, $i, j \in V$. We have that $A_{i,j} = 1$ if node i and node j are connected by an edge or $A_{i,j} = 0$ otherwise. We also have that A is symmetric (i.e., $A_{i,j} = A_{j,i}$) and we assume that no self-connections are present (i.e., $A_{i,i} = 0$).

A graph $G = (V, E)$ admits multiple possible modular organizations. A given organization partitions the node set V into a set of modules C and we define the number of modules as $t := |C|$. A module $c \in C$ is a node collection $V_c \subseteq V$ and we have that $\bigcup_{c \in C} V_c = V$ and $V_c \bigcap V_{c'} = \emptyset, c, c' \in C$ (modules have non-overlapping elements). We use notation $c(i) \in C, i \in V$ to denote the module that node $i \in \mathcal{V}$ belongs to. We define the binary module membership matrix $\delta \in \{0, 1\}^{n \times n}$ with entries:

$$\delta_{i,j} = \begin{cases} 1 & \text{if } c(i) = c(j) \\ 0 & \text{otherwise} \end{cases} \quad i, j \in V. \tag{2.2.1}$$

In other words, $\delta_{i,j} = 1$ if nodes i, j are in the same module or $\delta_{i,j} = 0$ otherwise. If all nodes are in the same module we have that $\sum_{i \in V} \sum_{j \in V} \delta_{i,j} = n \cdot n$; on the other hand, if all nodes are in separate modules, we have that $\sum_{i \in V} \sum_{j \in V} \delta_{i,j} = n$ because $\delta_{i,j} = 1$ for $i = j$ and $\delta_{i,j} = 0$ for $i \neq j$ (i.e., the membership matrix is the identity matrix). An important property of the membership matrix is that $\text{rank}(\delta) = t$. In other words, the number of modules equals the number of linearly independent columns (or rows) of δ. As can be seen, the membership matrix δ *encodes all relevant information associated with a given modular organization*.

To illustrate the relationship between the rank of the membership matrix and the number of modules, suppose that we have a graph with four nodes $V = \{a, b, c, d\}$ and $C = \{1, 2\}$ modules (and thus $t = 2$). Assume that nodes a and b are in the same module and nodes c and d are in the another module. Therefore, we have $\delta_{a,b} = \delta_{c,d} = 1$, and $\delta_{a,d} = \delta_{c,b} = 0$. We define the columns of δ corresponding to nodes a and c as δ_a and δ_c, and we would like to show that these columns are linearly independent. Equivalently, we want to prove that the only solution to the linear system

$$S_1 \delta_{a,b} + S_2 \delta_{c,b} = 0 \tag{2.2.2a}$$

$$S_1 \delta_{a,d} + S_2 \delta_{c,d} = 0 \tag{2.2.2b}$$

is $S_1 = S_2 = 0$. Upon substitution of $\delta_{i,j}$ in the above equations we obtain that $S_1 = S_2 = 0$ and we thus have $\text{rank}(\delta) = t = 2$.

The partitioning of the graph $G = (V, E)$ into modules induces an organization with a given connectivity inside modules (intra-modules) and between modules (inter-modules). We measure modularity of an organization as the density of internal module edges relative to the total number of edges. This measure is known as the *graph coverage* and can be computed as:

$$\text{cov}(\delta) = \frac{1}{m} \sum_{c \in C} |E(c)|$$

$$= \frac{1}{2m} \sum_{i,j \in V} \delta_{i,j} A_{i,j} \tag{2.2.3}$$

where $|E(c)|$ denotes the number of edges in module $c \in C$. The factor $1/2$ eliminates the repeated counting of edges in the adjacency matrix A.

We define the modularity measure of system G as:

$$M_t := \max_{\delta} \ \text{cov}(\delta) \tag{2.2.4a}$$

$$\text{s.t. rank}(\delta) \geq t. \tag{2.2.4b}$$

In other words, the modularity measure is computed by solving a rank-constrained optimization problem. We note that $\sum_{i,j \in V} \delta_{i,j} A_{i,j}$ is at least zero and at most $2m$ and thus $\text{cov}(\delta) \in [0, 1]$. The $\text{cov}(\delta) = 1$ case corresponds to $\sum_{i,j \in V} \delta_{i,j} A_{i,j} = 2m$ and occurs when $\delta_{i,j} = 1$ for all $i, j \in V$ (all nodes are in one module and thus $t = 1$). The $\text{cov}(\delta) = 0$ case corresponds to $\sum_{i,j \in V} \delta_{i,j} A_{i,j} = 0$ and occurs when $\delta_{i,j} = 0$ for all $i, j \in V$ such that $A_{i,j} = 1$. For the maximum possible rank $n = t = \text{rank}(\delta)$, this occurs when $\delta_{i,i} = 1$ and $\delta_{i,j} = 0$ (the membership matrix is the identity matrix). In general, we have that $\text{cov}(\delta)$ is large when connectivity between modules is sparse (connectivity inside modules is dense) and we have that $\text{cov}(\delta)$ is small when the connectivity between modules is dense (inside modules is sparse).

Consistent with the definition of graph coverage, the maximum possible value for M_t is achieved when all nodes are contained in one module ($t = 1$). Consequently, for connected graphs, the value of M_1 is always one. In this case, the modularity measure is given by the unconstrained problem $M_1 = \max_{\delta} \ \text{cov}(\delta)$. Restricting the number of modules to any value $t > 1$ forces placement of edges outside modules and thus decreases the modularity measure M_t. In the limit when we restrict $t = n$, we obtain the minimum modularity M_n (corresponding to the case in which every node is in a different module). The proposed modularity measure M_t thus naturally captures trade-offs between graph connectivity and number of modules.

The proposed definition of modularity is intuitive from a manufacturing perspective but alternative definitions exist in the literature (particularly in scientific applications). In the supplementary material we provide a perspective on alternative definitions along with their advantages and limitations.

2.2.2 *Illustrative Example*

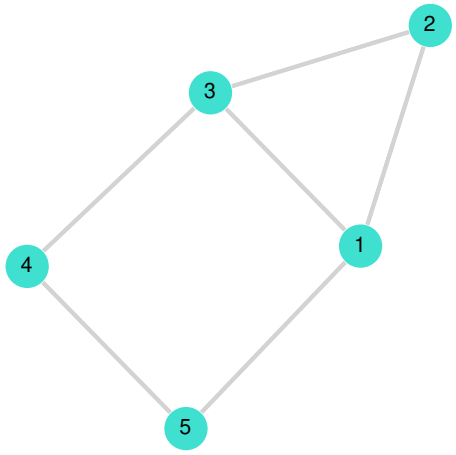


Figure 2.2: Example graph $G = (V, E)$ used to illustrate graph theoretical concepts.

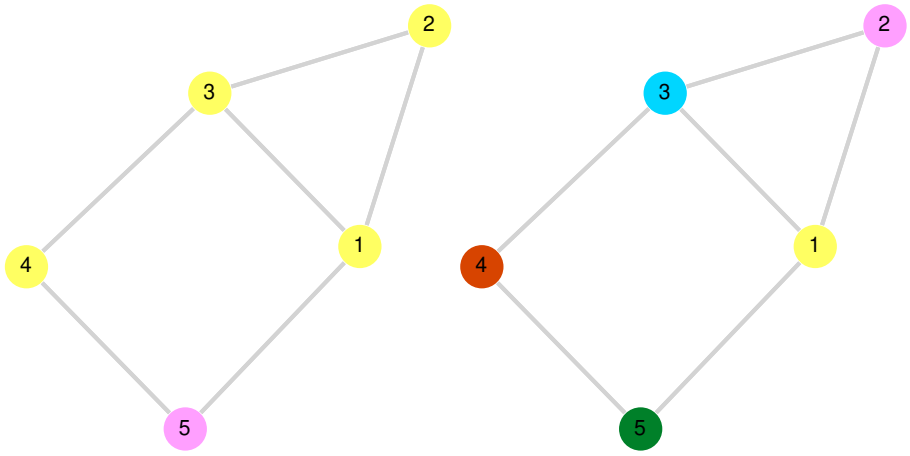


Figure 2.3: Modular organizations for example graph with $t = 2$ (left) and $t = 5$ (right).

We use a simple graph (see Figure 2.2) with $n = 5$ nodes and $m = 6$ edges to illustrate

the concepts. The adjacency matrix A of this graph is:

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

The membership matrices δ for two modular organizations with $t = 2$ and $t = 5$ (see Figure 2.3) are:

$$\delta = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \delta = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The matrix on the left has $\text{rank}(\delta) = t = 2$ and is a solution of problem (2.2.4) with $t = 2$ and modularity measure is $M_2 = 4/6$ (coverage $\text{cov}(\delta) = 4/6$). Because this is a solution to (2.2.4), any alternative configuration with $t = 2$ must have $M_2 \leq 4/6$. Upon inspection, one can indeed see that the number of internal module edges is four and the total number of edges is $m = 6$. The identity matrix on the right indicates that each node belongs to a module and thus $\text{rank}(\delta) = t = 5$ and is the solution of problem (2.2.4) with $t = 5$ and $M_5 = 0$ (coverage is $\text{cov}(\delta) = 0$).

2.2.3 *Optimization Formulations for the Modularity Measure*

We proceed to show that the proposed modularity measure can be formulated as a mixed-integer quadratic program. To motivate our discussion, we define the binary variable matrix $x \in \{0, 1\}^{n \times n}$ with entries $x_{i,j} := (1 - \delta_{i,j})$. The *unconstrained* modularity measure

can be computed by using the mixed-integer linear program (MILP):

$$\max_{x} \ \frac{1}{2m} \sum_{i,j} A_{i,j}(1 - x_{i,j}) \tag{2.2.5a}$$

$$\text{s.t. } x_{i,j} \leq x_{i,k} + x_{k,j}, \ j > i, i, j, k \in \mathcal{V} \tag{2.2.5b}$$

$$x_{i,i} = 0, i \in V \tag{2.2.5c}$$

$$x_{i,j} = x_{j,i}, i, j \in \mathcal{V}. \tag{2.2.5d}$$

The first constraint enforces the logic that, if i and j are in the same module and i and k are in the same module, then j and k must be in the same module. The second and third constraints capture basic logic that follows from the definition of the membership matrix. This formulation highlights intuitive connections between modularity and mixed-integer formulations. Unfortunately, this MILP formulation does not offer direct control on the number of modules (which is needed to compute measure M_t).

To obtain direct control on the number of modules, we propose a mixed-integer quadratic (MIQP) programming formulation. Here, we define a module (partition) set $C := \{1, ..., t\}$ with dimension $t \leq n$. We define a binary variable matrix $x \in \{0, 1\}^{n \times t}$ with entries $x_{i,k} = 1$ if node $i \in V$ is in module $k \in C$ and $x_{i,k} = 0$ otherwise. Importantly, under these definitions, we have that:

$$\delta_{i,j} = \sum_{k \in C} x_{i,k} x_{j,k}, \quad i, j \in V. \tag{2.2.6}$$

Because of this, the modularity measure M_t can be computed by using the MIQP:

$$\max_{x} \ \frac{1}{2m} \sum_{i,j \in V} A_{i,j} \sum_{k \in C} x_{i,k} x_{j,k} \tag{2.2.7a}$$

$$\text{s.t. } \sum_{k \in C} x_{i,k} = 1, \ i \in V \tag{2.2.7b}$$

$$\sum_{i \in V} x_{i,k} \geq 1, \ k \in C \tag{2.2.7c}$$

The first constraint enforces the logic that a node can only belong to one module while the second constraint ensures that at least one node is assigned to each module. The MIQP formulation is expected to be more computationally intensive than the MILP formulation but it captures the features needed (i.e., enforces the rank constraint). For simplicity in the discussion, we transform the MIQP into a minimization problem with objective $-\frac{1}{2m}\sum_{i,j\in V} A_{i,j}\sum_{k\in C} x_{i,k}x_{j,k}$.

2.2.4 *Convexification of MIQP Formulation*

We have found that the MIQP (in minimization form) can be cast as a convex MIQP, which is solvable by modern solvers. To see this, we define the variable vector $x_k = (x_{1,k}, x_{2,k}, ..., x_{n,k}),\ k \in C$ and note that we can rewrite the objective function as:

$$\begin{aligned}\frac{1}{2m}\sum_{i,j\in V} A_{i,j}\sum_{k\in C} x_{i,k}x_{j,k} &= \frac{1}{2m}\sum_{k\in C}\sum_{i,j\in V} x_{i,k}A_{i,j}x_{j,k} \\ &= \frac{1}{2m}\sum_{k\in C} x_k^T A x_k \\ &= \frac{1}{2m}\text{vec}(x)^T H \text{vec}(x) \end{aligned} \tag{2.2.8}$$

where H is a block-diagonal matrix of the form:

$$H = \begin{bmatrix} A & & \\ & \ddots & \\ & & A \end{bmatrix} \tag{2.2.9}$$

and $\text{vec}(x) = (x_1, x_2, ..., x_t)$. We note that H is indefinite because the adjacency matrix A is indefinite. However, we note that the entries of $\text{vec}(x)$ are all binary at any feasible solution and thus $\text{vec}(x)^T e = \text{vec}(x)^T \text{vec}(x)$ holds (e is a vector of ones of the same dimension as $\text{vec}(x)$). Consequently, we can write the objective function in the equivalent form:

$$\frac{1}{2m}\text{vec}(x)^T H \text{vec}(x) = \frac{1}{2m}(\text{vec}(x)^T (H + I\rho)\text{vec}(x) - \rho \text{vec}(x)^T e). \tag{2.2.10}$$

for any positive $\rho \in \mathbb{R}_+$ and where I is the identity matrix. This equivalence follows from:

$$\begin{aligned}\frac{1}{2m}\text{vec}(x)^T H\text{vec}(x) &= \frac{1}{2m}(\text{vec}(x)^T(H + I\rho)\text{vec}(x) - \rho\text{vec}(x)^T I\text{vec}(x)) \\ &= \frac{1}{2m}(\text{vec}(x)^T(H + I\rho)\text{vec}(x) - \rho\text{vec}(x)^T e) \\ &= \frac{1}{2m}(\text{vec}(x)^T H\text{vec}(x) + \rho\text{vec}(x)^T e - \rho\text{vec}(x)^T e) \\ &= \frac{1}{2m}\text{vec}(x)^T H\text{vec}(x). \end{aligned} \tag{2.2.11}$$

As a result, we can always make the coefficient matrix of the MIQP $(H + I\rho)$ positive definite without affecting the solution and thus make the problem solvable using state-of-the-art solvers. The most obvious choice for ρ would be to use the smallest eigenvalue of H (the smallest eigenvalue of A).

2.2.5 *Modeling Extensions and Other Applications*

Mixed-integer programming formulations offer flexibility to impose requirements that might be of interest from a manufacturing perspective. For instance, we consider the extended formulation:

$$\max_x \ \frac{1}{2m}\sum_{i,j\in V} A_{i,j}\sum_{k\in C} x_{i,k}x_{j,k} \tag{2.2.12a}$$

$$\text{s.t.} \sum_{k\in C} x_{i,k} = 1, \quad i \in V \tag{2.2.12b}$$

$$\sum_{i\in V} x_{i,k} \geq 1, \quad k \in C \tag{2.2.12c}$$

$$\underline{D}_k \leq \sum_{i\in\mathcal{V}} x_{i,k}D_i \leq \bar{D}_k, \quad k \in C \tag{2.2.12d}$$

Here, the last constraint imposes module feature constraints. The quantity $D_i \in \mathbb{R}_+$ denotes the feature of each node and $\underline{D}_k, \overline{D}_k \in \mathbb{R}_+$ are lower and upper bounds for the features. This constraint can be used to enforce different module features such as weight, height, and number of nodes in a module. For instance, the number of nodes in a module

can be controlled by using the constraint:

$$\underline{D}_k \leq \sum_{i \in \mathcal{V}} x_{i,k} \leq \bar{D}_k, \quad k \in C. \tag{2.2.13}$$

The proposed formulation can also be extended to impose logic constraints to force/prevent nodes from being in the same modules and can be extended to identify multiple organizations that lead to the same modularity measure (e.g., by using no-good cuts). In fact, modern mixed-integer solvers can compute all solutions that give the same optimal objective value.

We highlight that the proposed modularity measure and MIQP formulation can be used in other applications that go beyond manufacturing. For instance, these tools can be used to identify optimal configurations for control architectures and optimal decomposition strategies for optimization problems (Daoutidis et al., 2019; Allman et al., 2019). In this context, constraints on the number and size of modules can be used to create balanced configurations (e.g., to handle computational load balancing issues).

It is important to emphasize that modularity is directly associated to the partitioning of a graph and, as such, there is no analytical representation for such a measure. In other words, one needs to specify the partition of the graph first and then compute the modularity measure. Because of this, modularity needs to be expressed as an optimization that implicitly finds a partition of maximum modularity.

2.2.6 *Alternative Modularity Measures*

Broadly speaking, modularity is a graph measure that captures the density of internal edges in the modules relative to the total number of edges in a graph. From a mathematical standpoint, the measure can be defined in different forms and the actual selection is driven by the application at hand (Fortunato, 2010). In this section, we discuss an alternative modularity measure that is widely used in network analysis (in order to highlight advantages and disadvantages in the context of manufacturing).

In pioneering work, Newman proposed to measure the modularity of the graph by comparing the density of the internal module edges relative to those found in a random graph with similar properties (Newman, 2006). The argument behind this definition is that modularity originates naturally in real systems from non-random structures and thus a graph with high modularity should be the one that deviates as much as possible from a random graph.

To derive Newman's modularity measure, we define a random graph that has the same degree distribution as the system graph $G = (V, E)$. The random graph is such that the probability that a node is connected to another is uniform. Consequently, the probability that an edge starts from or ends at node $i \in V$ is $k_i/2m$. We define the probability matrix $P \in \mathbb{R}^{n \times n}$ with entries

$$P_{i,j} = \frac{k_i}{2m} \cdot \frac{k_j}{2m} = \frac{k_i k_j}{4m^2}, \quad i, j \in V \tag{2.2.14}$$

denoting the probability of finding a connection between node i and j. The expected value of the number of edges between nodes i and j is given by

$$F_{i,j} = 2m \cdot \frac{k_i k_j}{4m^2} = \frac{k_i k_j}{2m}, \quad i, j \in V \tag{2.2.15}$$

At the core of Newman's measure is the modularity matrix $B \in \mathbb{R}^{n \times n}$ with entries $B_{i,j} = A_{i,j} - F_{i,j},\ i, j \in V$ (the matrix is the difference of the graph adjacency matrix and the adjacency matrix of the random graph). We can establish that the modularity matrix is symmetric:

$$\begin{aligned} B_{i,j} &= A_{i,j} - \frac{k_i k_j}{2m} \\ &= A_{j,i} - \frac{k_i k_j}{2m} \\ &= B_{j,i}, i, j \in V. \end{aligned} \tag{2.2.16}$$

Moreover, we have that $\sum_{i \in V} B_{i,j} = \sum_{i \in V} B_{j,i} = 0$ for all $j \in V$ (the modularity matrix has

normalized columns and rows). Newman noticed that this property induces desirable properties of the eigenvalues and eigenvectors of the modularity matrix.

For a given organization (partition into set of modules C) we define the quality function $Q \in \mathbb{R}$ that measures the density of internal edges inside modules relative to the fraction induced by the associated random graph. In mathematical terms:

$$Q = \sum_{c \in C} \left(\frac{|E(c)|}{m} - \left(\frac{\sum_{i \in V} k_i}{2m} \right)^2 \right), \tag{2.2.17}$$

where $|E(c)|$ is the number of intra-cluster edges in module c ($|E(c)|/m$ is the coverage) and k_i is the degree of node i. The quality can be expressed in terms of the membership matrix δ as:

$$\begin{aligned} Q(\delta) &= \frac{1}{2m} \sum_{i,j \in V} B_{i,j} \delta_{i,j} \\ &= \frac{1}{2m} \sum_{i,j \in V} \left(A_{i,j} - F_{i,j} \right) \delta_{i,j} \\ &= \frac{1}{2m} \sum_{i,j \in V} \left(A_{i,j} - \frac{k_i k_j}{2m} \right) \delta_{i,j}. \end{aligned} \tag{2.2.18}$$

Newman's modularity measure is given by the modular organization that achieves the maximum quality function:

$$M := \max_{\delta} \; Q(\delta) \tag{2.2.19}$$

One can show that the quality function $Q(\delta)$ can take any value in the range $[-1/2, 1]$ and thus $M \in [-1/2, 1]$.

The measure of Newman does not assume a number of modules (as the proposed measure does). One can extend the definition to control the number of modules by using the rank-constrained formulation:

$$M_t := \max_{\delta} \; Q(\delta) \tag{2.2.20}$$

$$\text{s.t. } \text{rank}(\delta) \geq t. \tag{2.2.21}$$

Diverse MILP and MIQP formulations have been proposed to compute the unconstrained and rank-constrained variants of the modularity measure of Newman (Agarwal and Kempe, 2008; Xu et al., 2007). Interestingly, we note that one can compute this measure by using a MIQP that is similar to that proposed in our work. The MIQP formulation takes the form:

$$\max_{x} \ \frac{1}{2m} \sum_{i,j \in V} B_{i,j} \sum_{k \in C} x_{i,k} x_{j,k} \tag{2.2.22a}$$

$$\text{s.t. } \sum_{k \in C} x_{i,k} = 1, \ i \in V \tag{2.2.22b}$$

$$\sum_{i \in V} x_{i,k} \geq 1, k \in C. \tag{2.2.22c}$$

The objective function of this problem can also be expressed as $\frac{1}{2m}\text{vec}(x)^T H \text{vec}(x)$ where H is a block-diagonal matrix of the form:

$$H = \begin{bmatrix} B & & \\ & \ddots & \\ & & B \end{bmatrix} \tag{2.2.23}$$

The modularity matrix B is indefinite (Newman, 2006) (and thus H is indefinite) but we can use the same convexification procedure outlined previously to reformulate the MIQP into a concave QP. The proposed formulation is more intuitive and compact that the MIQP formulations that exist in the literature (Xu et al., 2007). Bench-marking the computational performance of different formulations is left as an interesting topic of future work.

The modularity measure of Newman can in principle be used to guide modular designs in manufacturing. The interpretation of this measure, however, is less intuitive from this perspective (e.g., it can take negative values and is a measure relative to a random graph). Moreover, this measure is not monotonic in the number of modules (as the

measure proposed in this work is). Because of this, the measure of Newman does not naturally minimize the number of modules. As a result, using this measure to perform comparisons between designs and systems is more complicated. We also highlight that the measure of Newman is degenerate (many configurations can give the same measure) and this can introduce significant ambiguity in the analysis. The proposed mixed-integer programming formulations provide a mechanism to explore and mitigate this degeneracy. Unfortunately, there are significant computational challenges to apply mixed-integer techniques in the analysis of large graphs (the community detection problem is known to be NP hard). A large number of heuristic techniques have been developed in the literature to handle large graphs

2.3 Case Study

We use the proposed modularity measure and MIQP formulation to identify modular configurations for a dimethyl-ethyl (DME) production process from methanol (Bhattacharyya et al., 2012). Methanol is an intermediate product during the production of DME from natural gas and thus small modular DME plants can provide a potential pathway to help recover billions of cubic feet of natural gas that are currently stranded and flared. The DME process is intuitively partitioned by practitioners into three functional subsystems: the feed and reactor section, the DME purification section, and the methanol separation and the recycle section. The process flow diagram (PFD) and the subsystems are shown in Figure 2.4. We created a block and graph representation for the process. To do so, we represent each equipment unit and junction as a node and each flow connection as an edge. The block and graph representations are shown in Figure 2.5. We use node dimension as a feature that affects the system modularity (i.e., this affects transportation). Labels and dimensions for the nodes are presented in Table 2.1.

We first computed the unconstrained modularity measure while ignoring rank (number of modules) and dimension constraints. As expected, the solution of this problem

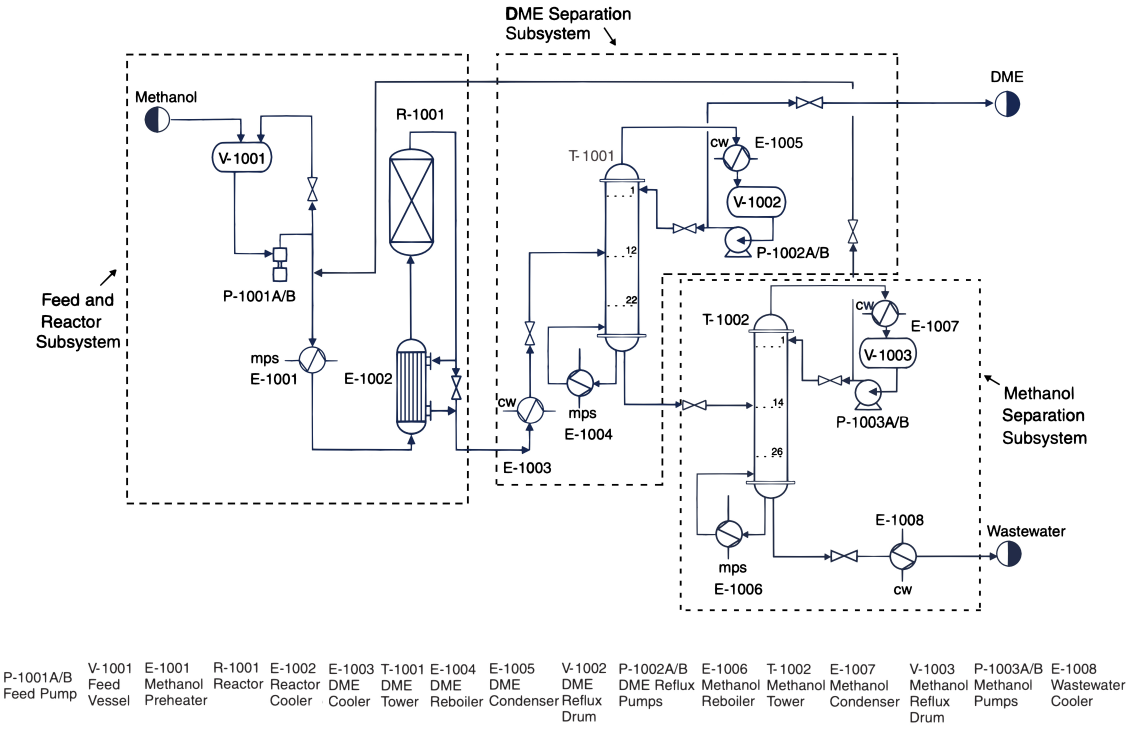


Figure 2.4: Flow diagram and subsystems for dimethyl-ethyl (DME) process (adapted from (Bhattacharyya et al., 2012)).

gives a modularity measure $M_1 = 1$ (i.e., all nodes are assigned to one module). We then computed the modularity measure by spanning the range $t \in [1, 6]$. The results are summarized in Table 2.2 and a visualization of each configuration is presented in Figure 2.6. We can see that, as the number of modules increases, the modularity measure decreases from $M_1 = 1$ to $M_6 = 0.875$. We thus have that, for a configuration with $t = 6$, 87.5% of the edges are inside the modules while 12.5% connect the modules (the configuration has sparse intermodule coupling).

For every value of t, we computed all possible equivalent configurations (solutions that give the maximum value of M_t). We do this in order to highlight that multiple configurations can give the same modularity measure. We found that the number of alternative solutions increases sharply with the increasing number of modules. This indicates that degeneracy increases with the number of modules and highlights the combinato-

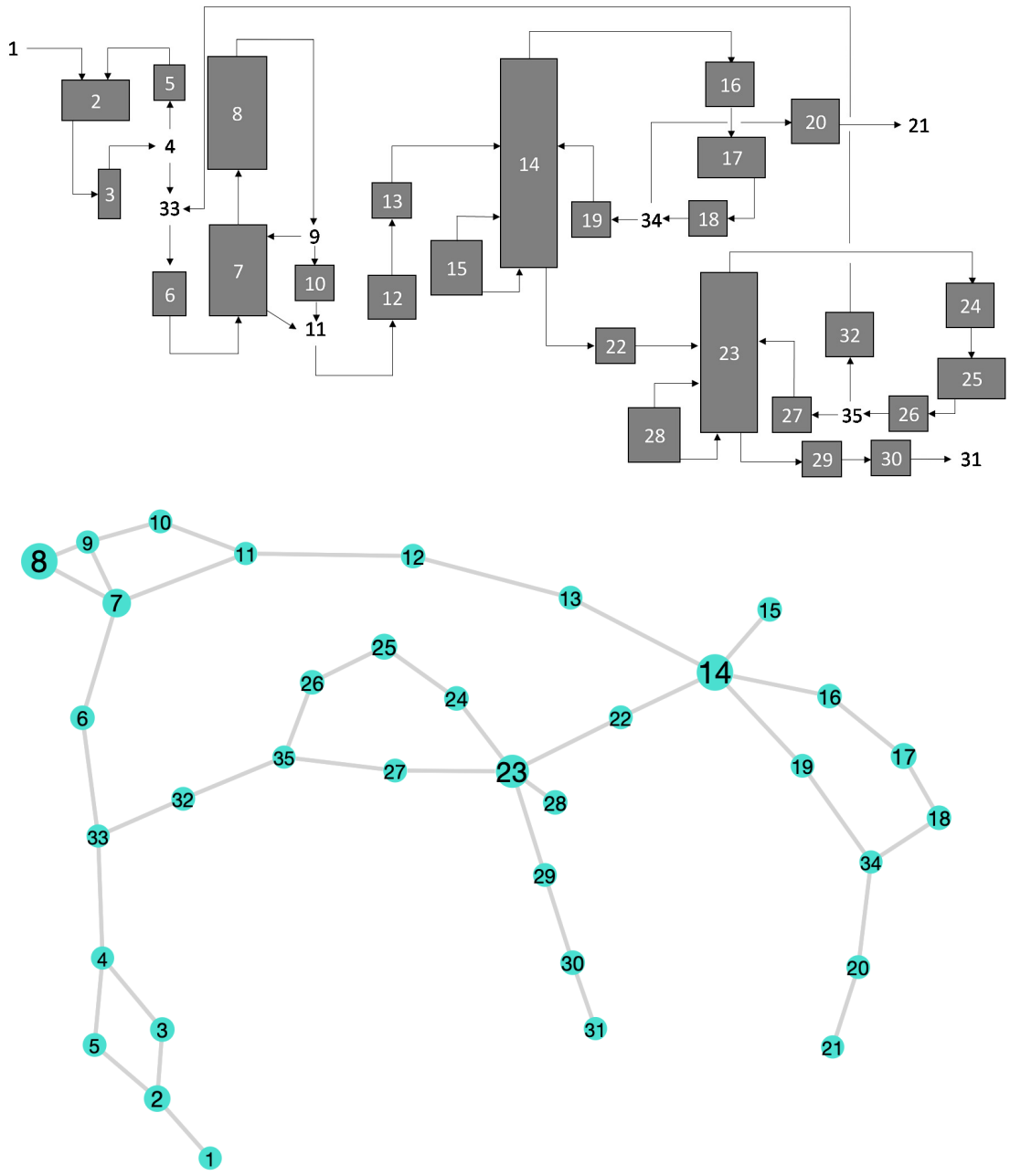


Figure 2.5: Block diagram representation (top) and graph representation (bottom) for DME process.

rial nature of the problem. This also indicates that there is significant flexibility to find configurations that satisfy additional requirements (such as dimension constraints).

We computed the modularity measure by considering dimension constraints (but ig-

Table 2.1: Node labels and dimensions for each node in the DME process.

Node	Equipment	Dimension	Node	Equipment	Dimension
1	Feed	0	19	Valve	1
2	V-1001	5	20	Valve	1
3	P-1001A/B	2	21	Product	0
4	Flow Junction	0	22	Valve	1
5	Valve	1	23	T-1002	15
6	E-1001	2	24	E-1007	2
7	E-1002	8	25	V-1003	4
8	R-1001	20	26	P-1003A/B	2
9	Flow Junction	0	27	Valve	1
10	Valve	1	28	E-1006	2
11	Flow Junction	0	29	Valve	1
12	E-1003	2	30	E-1008	2
13	Valve	1	31	Product	0
14	T-1001	20	32	Valve	1
15	E-1004	2	33	Flow Junc	0
16	E-1005	2	34	Flow Junc	0
17	V-1002	4	35	Flow Junc	0
18	P-1002A/B	2			

nore rank constraints). Here, we explore the impact of scaling-up the process and set the minimum dimension of each module to $\underline{D}_k = 20$ and the maximum dimension $\bar{D}_k = 40$. A visualization of the modular configuration is shown in Figure 2.7 and the associated node-module membership is shown in Table 2.3. Here, we also report the module dimensions $\sum_{i \in \mathcal{V}} x_{i,k} D_i$ for all $k \in C$. We observe that the dimension constraints induce an organization with $t = 3$ modules and the associated modularity measure is $M_3 = 0.925$ (only 7.5% of the edges connect modules). Interestingly, we can see that the resulting modu-

Table 2.2: Modularity measures and number of alternative configurations for DME process obtained with rank constraints (without dimension constraints).

Rank (t)	Modularity Measure (M_t)	# of Configurations	Solution Time (sec)
1	1	1	14.69
2	0.975	18	14.67
3	0.95	216	14.77
4	0.925	2016	15.68
5	0.9	15120	15.16
6	0.875	90720	15.10

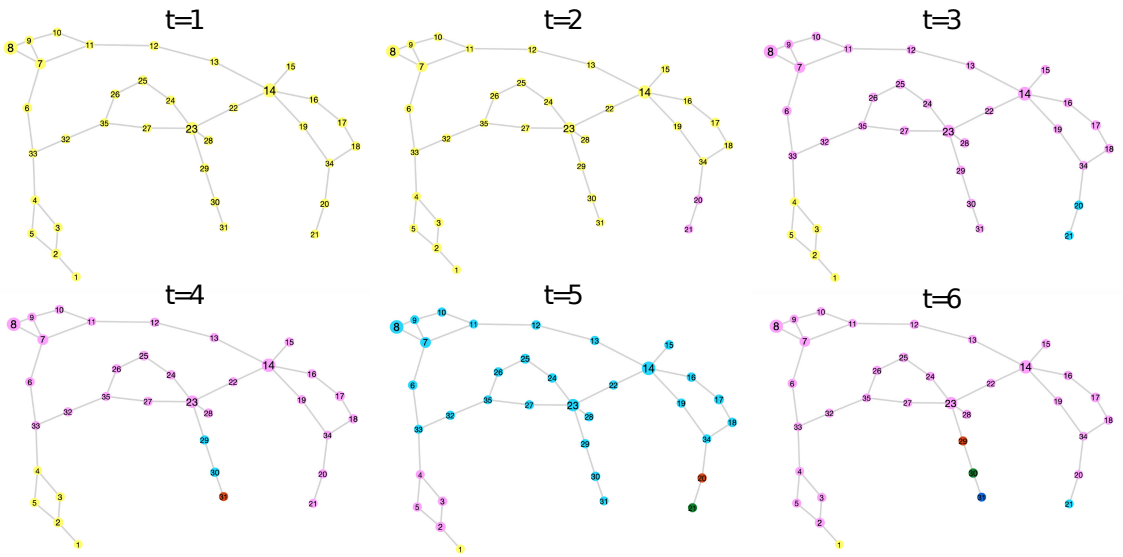
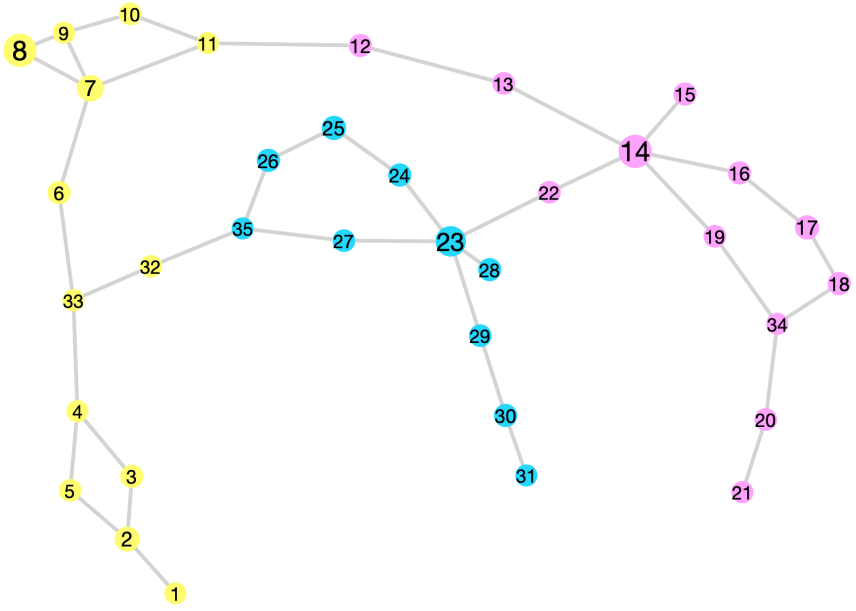


Figure 2.6: Sample modular configurations for DME process for $t \in [1, 6]$.

lar organization is the same as the functional organization shown in Figure 2.4 (with the exception of valves). In fact, we found that the modularity measure that result from the functional organization of Figure 2.4 is 0.925 and is thus optimal. This highlights that practitioners use natural logic to modularize systems and that the proposed modularity measure is intuitive.

We explored the effect of scaling up and down on the modularity measure by scaling



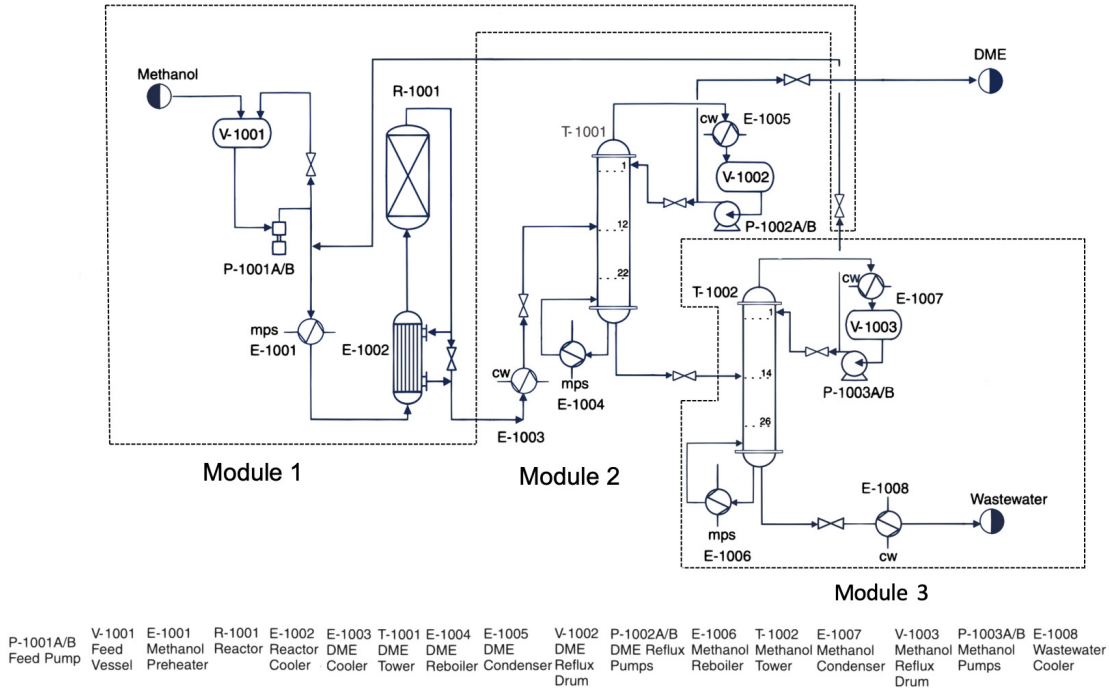


Figure 2.7: Modular configuration for baseline DME process obtained under dimension constraints (adapted from (Bhattacharyya et al., 2012)). Graph representation (top) and corresponding flowsheet (bottom).

Table 2.3: Modular configuration for DME process obtained under dimension constraints.

Module	Nodes	Module Dimension
1	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 32, 33]	40
2	[12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 34]	36
3	[23, 24, 25, 26, 27, 28, 29, 30, 31, 35]	29

the equipment unit dimensions. Scaling results are summarized in Table 2.4. We recall that the baseline measure value is $M_3 = 0.925$. As expected, we observe that the number of modules increases and the modularity measure decreases as we scale up the process. By scaling the equipment units up by 20% the measure decreases to $M_2 = 0.875$. The modularity measure achieves its ideal value of $M_1 = 1$ when the baseline process is scaled down by 30%. This highlights that the modularity measure proposed is consistent and that dimension constraints can also be used to implicitly control the number of modules.

Table 2.4: Effect of DME process scaling on modularity measure.

Scale	Measure (M_t)	Rank (t)	Solution Time (sec) with `Gurobi`	Solution Time (sec) with Convexification
0.3	1.0	1	15.59	16.91
0.5	0.95	2	15.25	16.17
1 (Baseline)	0.925	3	15.36	17.67
1.2	0.875	4	15.80	16.48
1.5	0.775	5	15.37	16.73
1.9	0.75	6	16.05	17.01

The MIQPs were solved using `Gurobi` (version 0.6.0) and were implemented in the Julia-based `JuMP` modeling framework. We use `GraphPlot` and `LightGraphs` for graph

manipulation and visualization. All node needed to reproduce the results can be found in https://github.com/zavalab/JuliaBox/tree/master/ModularityMeasures. We solved the MIQP problems by convexifying them directly. To do so, the minimum eigenvalue of the adjacency matrix A is -2.62 and thus we used $\rho = 3$. We highlight that `Gurobi` can also automatically convexify the problem (convexification by the user is not needed). We confirmed that both approaches give the same solutions (`Gurobi` gives slightly better times). The solution times obtained are in the range of 14 to 17 seconds (these are reported in Table 2.4).

Table 2.5: Comparison of the number of solutions for problems with and without dimension constraints (DCs).

Rank (t)	Measure (M_t) without DCs	# of Configurations without DCs	Measure (M_t) with DCs	# of Configurations with DCs
3	0.950	216	0.925	78
4	0.925	2016	0.875	384
5	0.900	15120	0.775	1920

We computed the number of solutions for problems with dimension constraints for $t = 3$, $t = 4$ and $t = 5$ and compared against the number of solutions obtained without dimension constraints. The results are summarized in Table 2.5. We can see that for all cases, the number of solutions are drastically reduced when dimensional constraints are added. This highlights the importance of enforcing additional module features to mitigate the natural degeneracy of modularity measures. In particular, other modularity measures used in the scientific literature, such as that of Newman, are degenerate (i.e., different organizations give the same modularity measure) and this degeneracy can introduce ambiguity in the analysis. Mixed-integer programming approaches allow us to systematically explore the set of feasible solutions.

2.4 Conclusions and Future Work

In this chapter, we propose a measure to quantify the modularity of industrial production (manufacturing) systems and optimization formulations to compute it. From a manufacturing perspective, we argue that a system is deemed modular if: i) the equipment units that comprise it form clusters (modules) of dense connectivity (i.e., difficult module assembly tasks are performed off-site), ii) connectivity between modules is sparse (i.e., easy assembly tasks are performed on-site), iii) the number of modules is small, and iv) the module dimensions facilitate transportation. We show that the measure proposed satisfies these requirements and that it can be computed by solving a convex mixed-integer quadratic program. Moreover, this formulation allows us to capture logical constraints associated with module dimensions and node-module membership restrictions. We provide a discussion on advantages and disadvantages of alternative modularity measures used in different scientific and engineering communities. Our results seek to highlight conceptual and computational challenges that arise from the need to define and quantify modularity in a manufacturing context.

As part of future work, we are interested in using the proposed measure to guide the synthesis of systems with desired modularity properties. To tackle scalability issues, we are also interested in exploring computational strategies to analyze large-scale graphs. Moreover, there exist interesting synergies of the modular design principles discussed in this work with modular design principles of control architectures (Daoutidis et al., 2019). This is because in both cases one is implicitly seeking to minimize the degree of interaction between modules. Using the proposed measure to understand the interplay between modular design and control is an interesting topic of future work.

3

BENEFITS OF MODULAR DESIGN - SPATIAL FLEXIBILITY

This chapter talks about the spatial flexibility brought by deploying modular technologies, and is published in Shao and Zavala (2019) titled "Space-Time Dynamics of Electricity Markets Incentivize Technology Decentralization".

3.1 Introduction

Decentralization of technologies for power generation (e.g., power plants), consumption (e.g., manufacturing facilities and data centers), and storage (e.g., batteries) is an ongoing industrial trend (Stankiewicz et al., 2000; Ramshaw, 1999). From the perspective of an independent system operator (ISO) of the power grid, decentralization is desirable as it can provide spatial flexibility to control network flows and to overcome limited transmission infrastructure (Buchholz, 2010; Kim et al., 2017a). In addition, large centralized power generation and consumption facilities can become liabilities during extreme weather or cyber attacks (Lier and Grünewald, 2011). To give an idea of the risk that large centralized facilities pose to the power grid, consider the fact that the load of a conventional ammonia manufacturing plant is around 64 MW (Egenhofer et al., 2017) and that the load of a large data center reaches 50 MW (Avgerinou et al., 2017) (equivalent to the load

of tens of thousands of homes). Similarly, the power supply of a large centralized power plant such as the Hammond plant in Georgia is 800 MW (Georgia Power). The growing demand from large data centers is of particular concern as it is projected that, within the next decade, the loads from such facilities will represent over 20% of the total grid load (Kim et al., 2017a). Another issue associated with centralized facilities is that they provide limited investment flexibility to mitigate long-term risks in electricity prices and policy. The need to mitigate investment risks is promoting the development and deployment of smaller-scale (modular) technologies (Guo et al., 2009; Wu et al., 2009; Palys et al., 2018; Kim et al., 2017b). On the other hand, it is well-known that large centralized systems benefit from economies of scale and thus a strong trade-off exists between expected profit and risk.

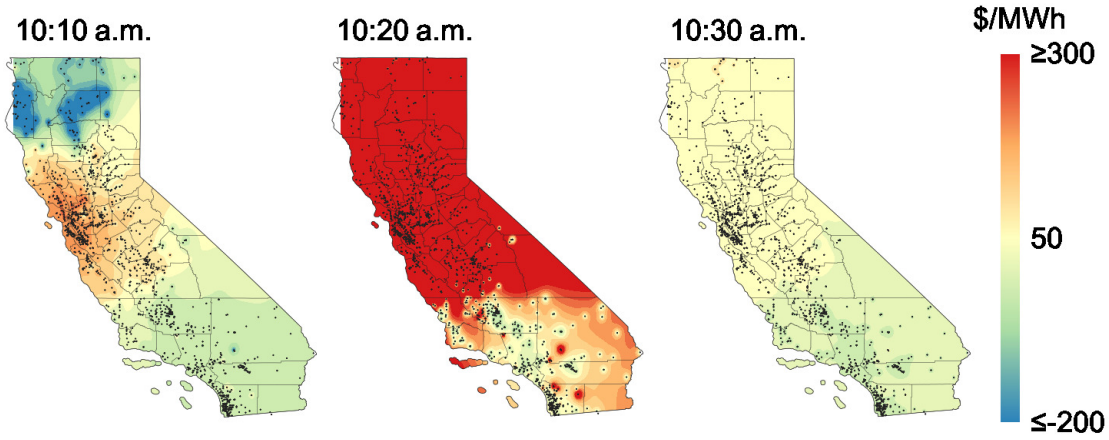


Figure 3.1: Electricity price fluctuation in the RTM on February 5, 2015 in CAISO.

Because electricity prices are a key driving factor in the revenue/cost of facilities, space-time price fluctuations must be considered in investment and operating decisions. For instance, power generation and consumption facilities often sell/purchase electricity in the Day-Ahead Energy Market (DAM) as opposed to the Real-Time Energy Market (RTM) to minimize risk, as the former is far less volatile (ISO New England; Dowling et al., 2017; Conejo et al., 2005). The growing share of renewable power in the supply portfolio is also introducing stronger market volatility and risk (Johnson and Oliver, 2016), as

unpredictable weather events can disrupt these renewable technologies and thus cause intermittent and volatile electricity supply. This issue is exacerbated by the lack of sufficient elastic (flexible) demand. The temporal and spatial volatility of electricity prices in RTM is illustrated in Figure 3.1. Here, we show the nodal price change over 20 minutes for a specific day in California. We see that, under a 20-minute period, the average electricity price increases from 48.42 USD/MWh to 592.33 USD/MWh and then drops to 35.15 USD/MWh. Here, we also see that such fluctuations are less abrupt at some network locations. Price volatility is less severe in day-ahead markets; in fact, day-ahead markets are precisely designed to pre-allocate generation and loads in the network in order to help participants mitigate profit risk (Zavala et al., 2017). On the other hand, the average RTM price is typically lower than the average DAM price. Consequently, there exists a premium to participate in the DAM (in order to avoid RTM volatility and associated risk). This suggests that there exists an economic incentive to decentralize (diversify) generation and load assets over multiple network locations in order to exploit spatial correlations in DAM and RTM prices (and with this avoid large premia). Similarly, spatial variations in DAM and RTM prices can be exploited by decentralized facilities to maximize profit. For instance, large cloud computing providers are currently placing data centers strategically in the network in order to avoid large electricity costs (Kim et al., 2017b). One could also envision that small modular manufacturing facilities can be relocated to exploit more favorable prices. A challenge that arises in this context is that DAM and RTM prices exhibit complex spatio-temporal dynamics and correlation patterns (Wang and Hobbs, 2014). As a result, it is *non-trivial* to identify suitable degrees of asset decentralization and optimal locations for such assets.

In this work, we propose a computational framework for analyzing economic incentives created by space-time dynamics of electricity markets. Our framework is based on an asset placement formulation that seeks to find optimal locations for generation and load (consumption) assets in the network that minimize profit risk. We show that an unconstrained version of this problem can be cast as an eigenvalue problem. Under

this representation, optimal network allocations are eigenvectors of the space-time price covariance matrix, while the eigenvalues are the profit variances that result from such allocations. Consequently, risk analysis can be performed in a systematic and computationally efficient manner by using principal component analysis (PCA). We construct a constrained placement problem that captures constraints on the types of assets and that trade-offs risk and expected profit. Unfortunately, for the ISO-scale data sets of interest, this problem is a large-scale mixed-integer quadratic programming (MIQP) problem that cannot be solved with current solvers. We use the mean absolute deviation as an alternative risk measure to obtain a more scalable (but still challenging) mixed-integer linear program (MILP). Analysis using the California ISO (CAISO) market data for 2015 reveals that space-time market dynamics provide significant incentives for strategic diversification and asset placement but that complete mitigation of revenue risk is only possible by *simultaneous* investment in decentralized generation and load assets (which can also be achieved by using batteries or hybrid systems such as microgrids or other prosumers). These results are of relevance given the recent interest in the deployment of small-scale modular technologies. We highlight that our work focuses on the use of real (but historical) data to conduct analysis; as such, the study is realistic but has limited predicted power. Unfortunately, existing forecasting techniques for economic time-series data focus on uni-dimensional data (Ledoit and Wolf, 2004), while the market data set considered here is high-dimensional (reaching thousands of locations that are correlated in space and time). As part of future work, we will investigate forecasting strategies for such high-dimensional data sets.

In the following sections, we motivate our discussion by conducting a basic space-time analysis of electricity markets in California. In section 3.2, we formulate the electricity market clearing model and explore the effect of modularization on social welfare and electricity prices. In Section 3.3 we formulate the technology placement problem, interpret it as an eigenvalue problem, and provide scalable constrained variants. A detailed analysis of the California ISO data set using the placement problem formulations

is provided in Section 3.4.

3.2 Electricity Market Clearing Model

In this section, we formulate the electricity market clearing model and use optimization tools to find out the effect of modularization on social welfare and electricity prices.

3.2.1 *Market Setting*

We consider a deterministic market setting that comprises a set of nodes $\mathcal{N}$, transmission lines that between each node $\mathcal{F}$, power suppliers $\mathcal{S}$, and power consumers $\mathcal{D}$. Each node $n \in \mathcal{N}$ is associated with a power supplier, a power consumer and power flows in and out of that node. At a single node, each power supplier $i \in \mathcal{S}$ has its own capacity $\bar{s}_i \in \mathbb{R}_+$ and bidding price $\alpha_i^s \in \mathbb{R}_+$, and each power consumer $j \in \mathcal{D}$ has its own capacity $\bar{d}_j \in \mathbb{R}_+$ and bidding price $\alpha_j^d \in \mathbb{R}_+$. The transmission line $\ell \in \mathcal{F}$ has a two-directional flow $f_\ell \in \mathbb{R}$, and satisfies $0 \leq f_\ell \leq \bar{f}_\ell$. A positive value of f_ℓ corresponds to a flow into the node, while a negative value of f_ℓ corresponds to a flow out of the node. We also define a market clearing price (i.e., locational marginal price) for each node $n \in \mathcal{N}$ as π_n. A shorthand notation s, d, f is used to refer to all supply, demand and flow in the electricity market.

3.2.2 *Node Balance*

Each node $n \in \mathcal{N}$ in the network should satisfy the following flow balance:

$$\left(\sum_{i \in \mathcal{S}_n} s_i - \sum_{j \in \mathcal{D}_n} d_j\right) + \left(\sum_{\ell \in \mathcal{F}_n^{out}} f_\ell - \sum_{\ell \in \mathcal{F}_n^{in}} f_\ell\right) = 0, \quad n \in \mathcal{N} \tag{3.2.1}$$

The terms inside the first parenthesis are the supply and demand of electricity at each node. The terms inside the second parenthesis are the total inflows and outflows of a node. The difference in the supply and demand at each node is balanced by the net flow

that comes in or flows out of the node.

3.2.3 *Capacity Constraints*

The supply, demand, and flows of each node are bounded by:

$$0 \leq s_i \leq \bar{s}_i,\ i \in \mathcal{S} \tag{3.2.2a}$$

$$0 \leq d_j \leq \bar{d}_j,\ j \in \mathcal{D} \tag{3.2.2b}$$

$$0 \leq f_\ell \leq \bar{f}_\ell,\ \ell \in \mathcal{F}. \tag{3.2.2c}$$

The upper bounds of the supply and the demand are the amount of electricity that each supplier and consumer report to the ISO before it clears the market. The flow constraints are determined based on the quality of transmission lines.

3.2.4 *Market Clearing Formulation*

Every time the ISO clears the market, the goal is to maximize the social welfare, which includes any gain or loss on the transmission lines and the difference between the benefit of all the consumers and the overall cost of all the suppliers. The social welfare function in a coordinated market that captures the supply, the demand, and the transmission lines is:

$$\varphi = \sum_{j \in \mathcal{D}} \alpha_j^d d_j - \sum_{i \in \mathcal{S}} \alpha_i^s s_i - \sum_{\ell \in \mathcal{F}} \alpha_\ell^f f_\ell \tag{3.2.3}$$

Therefore, the optimization problem that the ISO faces is to maximize the social welfare function. The market clearing problem becomes:

$$\max \sum_{j \in \mathcal{D}} \alpha_j^d d_j - \sum_{i \in \mathcal{S}} \alpha_i^s s_i - \sum_{\ell \in \mathcal{F}} \alpha_\ell^f f_\ell \tag{3.2.4a}$$

$$\text{s.t.} \left(\sum_{i \in \mathcal{S}_n} s_i - \sum_{j \in \mathcal{D}_n} d_j \right) + \left(\sum_{\ell \in \mathcal{F}_n^{out}} f_\ell - \sum_{\ell \in \mathcal{F}_n^{in}} f_\ell \right) = 0, \quad n \in \mathcal{N} \tag{3.2.4b}$$

$$0 \leq s_i \leq \bar{s}_i, \ i \in \mathcal{S} \tag{3.2.4c}$$

$$0 \leq d_j \leq \bar{d}_j, \ j \in \mathcal{D} \tag{3.2.4d}$$

$$0 \leq f_\ell \leq \bar{f}_\ell, \ \ell \in \mathcal{F}. \tag{3.2.4e}$$

3.2.5 *Electricity Price and Market Player Profits*

The dual variable of the above optimization model represents the market clearing price at node n for $n \in \mathcal{N}$ as π_n, and the profit of each market player is represented as follows:

$$\phi_i^s(\pi_i, s_i) := (\pi_i - \alpha_i^s)s_i, \ i \in \mathcal{S} \tag{3.2.5a}$$

$$\phi_j^d(\pi_j, d_j) := (\alpha_j^d - \pi_j)d_j, \ j \in \mathcal{D} \tag{3.2.5b}$$

$$\phi_\ell^f(\Delta\pi_\ell, f_\ell) := (\Delta\pi_\ell - \alpha_\ell^f)f_\ell, \ \ell \in \mathcal{F}. \tag{3.2.5c}$$

Equation 3.2.5a shows the profit of each supplier in the market. $\pi_i s_i$ is the revenue gained by supplier $i \in \mathcal{S}$ from selling electricity, and $\alpha_i^s s_i$ represents the cost of power production by the same supplier. Similarly, Equation 3.2.5b shows the profit each consumer achieves in the market. $\pi_j d_j$ is the cost of consumer $j \in \mathcal{D}$ by paying the electricity, and $\alpha_j^d d_j$ is the revenue gained by the consumer, which can be thought of as the price that the customer values the electricity. Equation 3.2.5c shows the profit gathered by the transmission line operators. Here, $\Delta\pi_\ell$ represents the difference in the electricity price between the two nodes that the transmission line connects. $\Delta\pi_\ell f_\ell$ represents the revenue gained and $\alpha_\ell^f f_\ell$ represents the cost of the transmission line.

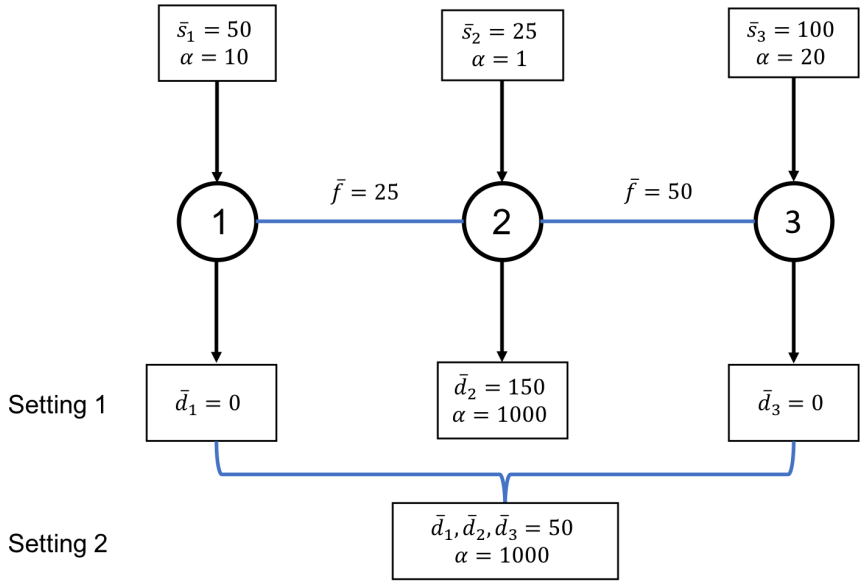


Figure 3.2: Scheme of system 1 with 2 settings

3.2.6 *Case Study*

Now that we have established the fundamentals of the electricity market, we use a simple setting, in which observations can be made comparatively easily, to explore the effect of modularization on social welfare and electricity prices. We seek to make a generalized conclusion based on our observations, and prove certain properties mathematically in the following sections.

Figure 3.2 shows the same electricity power grid with two different market settings that we will use to demonstrate the idea of modularization. In the first setting, the system has three nodes, each of which is associated with a supply and a demand. Each supply and demand are deterministic, and they are independent of each other. Node 1 has a supply capacity of 50 MWh and a demand capacity of 0 MWh. Node 2 has a supply capacity of 25 MWh and a demand capacity of 150 MWh. Node 3 has a supply capacity of 100 MWh and a demand capacity of 0 MWh. For all three nodes, the bid prices for the suppliers, α_i^s, are {10,1,20} \$/MWh, while the bid prices for the consumers, α_i^d, are {1000,1000,10000} \$/MWh. The electricity can flow from node 1 to node 2, and from node

Table 3.1: System 1. Comparison of quantities, prices and social welfare

settings	d_j (MWh)	s_i (MWh)	π_i ($)	f_ℓ (MWh)	φ ($)
setting 1	{0,150,0}	{25,25,50}	{10,1000,20}	{25,-50}	98,650
setting 2	{50,50,50}	{50,25,75}	{20,21,20}	{0,-25}	147,950

Table 3.2: System 1. Comparison of the profits

settings	$\phi_i^s(\pi_i, s_i)$ ($)	$\phi_j^d(\pi_j, d_j)$ ($)	$\phi_\ell^f(\Delta\pi_\ell, f_\ell)$ ($)	Total Profit ($)
setting 1	24,975	0	73,675	98,650
setting 2	1,000	146,950	0	147,950

2 to node 3, both forward and backward but not from node 1 to node 3. The transmission capacities of both lines are also deterministic, and set to {25,50} MWh. The costs of transmission for both lines are set to be 1 $/MWh. In this setting, the electricity consumer at node 2 represents one of the chemical plants or data centers mentioned previously, which consumes a huge amount of electricity that may have a profound impact on power transmissions and electricity prices at different locations.

In the second setting, we still have 3 nodes, and all the settings remain the same except that now, the single demand at node 2 in setting 1 is replaced by 3 identical demand at all 3 nodes. That is, the demand capacity for node 1, 2 and 3 is all 50 MWh. In other words, instead of a single centralized chemical plant or data center, now it is decentralized to 3 modular facilities that each has the same electricity consumption and adds up to the amount that the centralized facility in setting 1 consumes. With the decentralized loads, we expect that the tension in the transmission lines can be relaxed, and the electricity can be transformed more freely among the three nodes. The described setups and the results for both settings are summarized in Table 3.1 and 3.2.

We can see that when we decentralize the electricity consumption among all three

nodes in setting 2, the social welfare improves by 50%, from 98,650$ to 147,950$. Both transmission lines changes from fully loaded in setting 1 to partially loaded or fully relaxed in setting 2. In setting 1, only 100 MWh of 150 MWh electricity demand is satisfied due to the limits of the transmission line. However, in setting 2, the relatively small energy consumption at each node can first be satisfied by the electricity supply at its own location, which relaxes the tension in the transmission lines, and creates the spatial flexibility that allows the ISO to satisfy more electricity demand elsewhere. Another characteristic of setting 2 is that the electricity prices at all 3 nodes are homogenized. In setting 1, again because of the limits in the transmission line and the shortage of electricity, the electricity price at node 2 is the same as the bid price for consumers, leaving no profits for them in the electricity market. However, when the tension in the transmission lines is relaxed, and all demands are satisfied in setting 2, the electricity market switches from a seller's market to a buyer's market. The spatial flexibility allows the ISO to allocate the electricity freely around the three nodes, homogenizing the prices at all locations. The absurdly high price at node 2 disappears, and profits for electricity consumers increase. Also shown in Table 3.2, the profits for the supplier and the transmission lines decrease, while the profit for electricity consumers increases. Most importantly, the total profit increases and the value matches the maximized social welfare. The profit change is mainly the result of homogenized electricity prices.

3.3 Optimal Placement Problem

In this section, we derive different variants of the optimal placement problem that will allow us to explore incentives provides by space-time dynamics of electricity markets.

3.3.1 *Unconstrained Formulation*

We capture the space-time price data in a matrix $\Pi \in \mathbb{R}^{m \times n}$. Here, the number of columns n is the number of spatial network locations (nodes) and the number of rows m is the

number of time points. The matrix entry $\Pi_{i,j}$ is interpreted as price at time i and at node j. We use $p_i := \Pi_{i,:} \in \mathbb{R}^n$, $i = 1, ..., m$ to denote all node prices at time i. We denote the set of spatial locations as $\mathcal{N} := \{1, ..., n\}$ and the set of all time realizations as $\mathcal{M} := \{1, ..., m\}$. All prices have units of USD/MWh and we construct separate matrices for DAM and RTM.

Given the space-time price data, we seek to identify optimal locations for loads and generators in the network that minimize the temporal profit variance (variance is used as a standard measure of risk and can also be interpreted as profit volatility). We define a node allocation vector $w \in \mathbb{R}^n$ and the profit function at time i as $\varphi(w, p_i) := \sum_{j \in \mathcal{N}} w_j \Pi_{i,j}$. We interpret a positive node allocation $w_j > 0$ as an injection of power (a generation asset incurring revenue for a positive price) and a negative node allocation $w_j < 0$ as a withdrawal of power (a load asset incurring cost for a positive price). The node allocations w_j have units of MWh. If the prices are negative, a positive allocation incurs cost and a negative allocation incurs a revenue. In other words, installing generators maximizes revenue, but we will see that the simultaneous installation of generators and loads is needed to minimize risk.

The temporal average of the profit is given by:

$$\mu_\varphi(w) = \frac{1}{m} \sum_{i \in \mathcal{M}} \varphi(w, p_i) \tag{3.3.6}$$

and the temporal variance is

$$\Sigma_\varphi(w) = \frac{1}{m-1} \sum_{i \in \mathcal{M}} (\varphi(w, p_i) - \mu_\varphi(w))^2. \tag{3.3.7}$$

The optimal placement problem consists of finding the allocation vector w that minimizes the profit risk. This problem is stated as:

$$\min_w \ \Sigma_\varphi(w). \tag{3.3.8}$$

We assume that the optimal allocation vector (denoted as w_1^*) satisfies the constraint $\|w_1^*\|_2 = 1$ (it is a vector of unit length), where $\|\cdot\|_2$ denotes the Euclidean norm. This constraint is interpreted as the distribution of a finite amount of power among the network nodes. We note that the placement problem is scale-invariant. In other words, replacing $w \to \gamma w$ for some $\gamma > 0$ in the optimization problem yields the same optimal allocations. This is because $\Sigma_\varphi(\gamma w) = \gamma^2 w^T S w$ (resulting in a linear scaling of the objective function). Consequently, imposing a constraint of the form $\|w\|_2 = 1/\gamma$ will yield the same optimal allocation obtained with the constraint $\|w\|_2 = 1$. This formulation seeks to exploit the space-time dynamics of the prices to identify node allocations for generation or load that minimize risk. This is a large-scale and continuous *quadratic program* (QP).

3.3.2 *Eigenvalue Interpretation*

An interesting observation that we make is that, under the special case with no temporal price correlations, the optimal placement problem described above can be interpreted as an *eigenvalue problem*. This connection reveals some interesting properties of the market prices. In the absence of temporal price correlations, the price at the spatial location (network node) j can be modeled as a random variable (denoted as P_j) and we use $P = \{P_1, ..., P_m\}$ to denote a random vector containing all node prices. Consequently, the matrix entry $\Pi_{i,j}$ is interpreted as the i-th time realization of the price P_j and we assume that the probability of the realization is $1/m$. In this case, p_i denotes the i-th realization of the spatial price vector P. Under this setting, the sample average of the profit approximates the expected value of the profit:

$$\mu_\varphi(w) \approx \mathbb{E}[\varphi(w, P)] \tag{3.3.9}$$

and the sample variance approximates the variance:

$$\Sigma_\varphi(w) \approx \mathbb{V}[\varphi(w, P)]. \tag{3.3.10}$$

Here, we recall that $\mathbb{V}[\varphi(w,P)] = \mathbb{E}[\varphi(w,P)^2] - \mathbb{E}[\varphi(w,P)]^2$. The key observation is that the profit variance is related to the price covariance as $\mathbb{V}[\varphi(w,P)] = w^T\mathbb{E}[(P-\mathbb{E}[P])(P-\mathbb{E}[P])^T]w$. This result can be from the following series of implications:

$$\begin{aligned}\mathbb{V}[\varphi(w,P)] &= \mathbb{E}[\varphi(w,P)^2] - \mathbb{E}[\varphi(w,P)]^2 \\ &= \sum_{j\in\mathcal{N}}\sum_{k\in\mathcal{N}} w_j w_k \mathbb{E}[P_j P_k] - \sum_{j\in\mathcal{N}}\sum_{k\in\mathcal{N}} w_j w_k \mathbb{E}[P_j]\mathbb{E}[P_k] \\ &= \sum_{j\in\mathcal{N}}\sum_{k\in\mathcal{N}} w_j w_k \mathrm{Cov}(P_j, P_k) \\ &= w^T\mathbb{E}[(P-\mathbb{E}[P])(P-\mathbb{E}[P])^T]w. \end{aligned} \tag{3.3.11}$$

One can derive a similar relationship between the sample profit and covariance matrix to establish $\Sigma_\varphi(w) = w^T\Sigma w$. Consequently, the optimal placement problem (3.3.8) can also be written as:

$$\min_w \; w^T\Sigma w \text{ s.t. } \|w\|_2 = 1. \tag{3.3.12}$$

This reveals that the placement problem is an *eigenvalue problem*. Accordingly, the optimal allocation vector w_1^* is the eigenvector corresponding to the minimum eigenvalue λ_1^* of the price covariance matrix Σ. Moreover, the minimum eigenvalue is the minimum profit variance ($\lambda_1^* = \Sigma_\varphi(w_1^*)$). The eigenvalue problem is also a QP but this can also be solved efficiently using standard techniques (e.g., QR or SVD).

The eigenvalue problem is the basis of principal component analysis (PCA). The first principal component is given by $(w_1^*)^T p_i$, $i \in \mathcal{N}$. In PCA, one extracts the entire eigenvalue spectrum of the price matrix to obtain all the principal components. For instance, to obtain the second smallest eigenvalue and corresponding eigenvector we add the linear orthogonality constraint $w^T w_1^* = 0$ to the eigenvalue problem (3.3.12). The solution of the new problem yields the eigenvector w_2^* and corresponding eigenvalue $\lambda_2^* = \Sigma_\varphi(w_2^*)$. Since adding the orthogonality constraint restricts the feasible space, we

have that $\Sigma_\varphi(w_2^*) \geq \Sigma_\varphi(w_1^*)$. This procedure is repeated to obtain the entire set of eigenpairs w_j^*, λ_j^*, $j \in \mathcal{N}$, where $\lambda_n^* = \Sigma_\varphi(w_n^*)$ is the maximum possible cost variance (obtained with the loading allocation w_n^*). In our context, this procedure provides useful information because it allows us to obtain a family of allocations w_j^*, $j \in \mathcal{N}$ and to rank them according to their profit variance. The eigenvectors can also be used to form a matrix W that can be used to project any price realization p_i into the space of the principal components as Wp_i. The projection can be used to identify clusters and/or outliers in the price data by analyzing only a subset of principal components.

3.3.3 *Constrained Formulation*

While mitigating profit variance is an important investment goal, obtaining a maximum expected profit is also important. Moreover, one often has constraints on the nature and capacity of assets that can be installed. We thus extend the placement problem (3.3.8) to capture these features. We impose an ℓ_1-norm constraint on the allocation vector w so that the total amount of power allocated adds up to one MWh and we add a condition that only one type of asset is allowed to be built at one location (either generation or load). Consequently, we can decompose the node allocation w_j into a generation $0 \leq w_{j,l} \leq 1$ and a load component $-1 \leq w_{j,g} \leq 0$ (which are mutually exclusive). This gives the following conflict resolution (multi-objective optimization) problem:

$$\max_{w} \ \{\mu_\varphi(w), -\Sigma_\varphi(w)\} \tag{3.3.13a}$$

$$\text{s.t.} \sum_{j\in\mathcal{N}} \left(|w_{j,l}|+|w_{j,g}|\right) = 1 \tag{3.3.13b}$$

$$0 \leq w_{j,g} \leq z_{j,g},\ j \in \mathcal{N} \tag{3.3.13c}$$

$$-z_{j,l} \leq w_{j,l} \leq 0,\ j \in \mathcal{N} \tag{3.3.13d}$$

$$z_{j,l} + z_{j,g} \leq 1,\ j \in \mathcal{N} \tag{3.3.13e}$$

$$z_{j,l}, z_{j,g} \in \{0,1\}, j \in \mathcal{N}. \tag{3.3.13f}$$

where $z_{j,l}$ and $z_{j,g}$ are binary variables that indicate if either a load or generation asset is installed at a particular location j. The constraint $z_{j,l} + z_{j,g} \leq 1$ indicates that either a load or a generator (but not both) can be installed at one location. Consequently, we have that $\sum_{j\in\mathcal{N}} \left(|w_{j,l}|+|w_{j,g}|\right) = \sum_{j\in\mathcal{N}}|w_j| = \|w\|_1$. This constraint is used to avoid degeneracy of the solution (e.g., adding a load and a generator in a given node has the same net effect as installing one generator or load). The objective function captures the trade-off between expected profit and risk (which are often conflicting). The use of binary variables allows us to enforce a sharp separation between loads and generators (a continuous formulation does not allow for this). This facilitates interpretability of the solution. Specifically, we aim to use the placement formulation to explore how space-time price dynamics provide incentives to install decentralize facilities for loads and generators.

Unfortunately, the constrained placement problem is a large-scale mixed-integer QP. This problem is intractable for the ISO-scale data sets considered in this work. Motivated by this limitation, we consider the mean absolute deviation as a risk measure. This is given by:

$$\mathrm{MD}(w) = \frac{1}{m}\sum_{i\in\mathcal{M}} |\varphi(w, p_i) - \mu_\varphi(w)| \approx \mathbb{E}[|\varphi(w, P) - \mu_\varphi(w)|]. \tag{3.3.14}$$

This risk measure is used to formulate the placement problem:

$$\max_{w} \ \{\mu_\varphi(w), -\mathrm{MD}(w)\} \tag{3.3.15a}$$

$$\text{s.t.} \sum_{j\in\mathcal{N}} \left(|w_{j,l}|+|w_{j,g}|\right) = 1 \tag{3.3.15b}$$

$$0 \leq w_{j,g} \leq z_{j,g},\ j \in \mathcal{N} \tag{3.3.15c}$$

$$-z_{j,l} \leq w_{j,l} \leq 0,\ j \in \mathcal{N} \tag{3.3.15d}$$

$$z_{j,l} + z_{j,g} \leq 1,\ j \in \mathcal{N} \tag{3.3.15e}$$

$$z_{j,l}, z_{j,g} \in \{0, 1\}, j \in \mathcal{N} \tag{3.3.15f}$$

which can be cast as a mixed-integer linear program that is still large-scale scale but tractable with existing tools. The constrained placement problem is also scale-invariant. In other words, replacing $w \rightarrow \gamma w$ for some $\gamma > 0$ yields the same optimal allocations. This is because $MD(\gamma w) = \gamma MD(w)$, and $\mu_\varphi(\gamma w) = \gamma \mu_\varphi(w)$ (resulting in a linear scaling of the objective function). Consequently, imposing a constraint of the form $\|w\|_1 = 1/\gamma$ will yield the same allocation obtained with the unit-length constraint $\|w\|_1 = 1$. The constraints set capacity of load and generators, and the constraint on the binary variables ensures that only one type of technology is allowed at each location. We can use the above formulation to understand the impacts of installing only certain types of assets or at certain locations. For instance, if we only wish to install generation assets, we set all $z_{j,l}$ to zero.

3.4 Results and Discussion

In this section we use conduct a basic statistical analysis for an electricity market data set of CAISO and use the optimal placement formulation to analyze economic incentives created by the DAM and RTM.

3.4.1 *Volatility Analysis of Electricity Markets*

In the DAM, electricity prices are updated hourly and market participants commit to buy or sell power one day before real-time operation, thus avoiding price volatility. This market produces one financial settlement per day. In the RTM market, prices are updated every 5 minutes and participants commit to buy or sell electricity over the course of the operating day. This market seeks to balance discrepancies between the day-ahead commitments and the actual real-time generation and loads seen in the power grid (e.g., due to unexpected variations in renewable power supply, equipment failures, and so on) . The DAM and RTM work together to produce a multi-settlement system that balances power at different timescales and at thousands of network locations (Dowling et al., 2017).

Usually, electricity prices in the DAM are usually less volatile but are on average higher than RTM prices, and thus market participants can participate strategically in either or both of these markets.

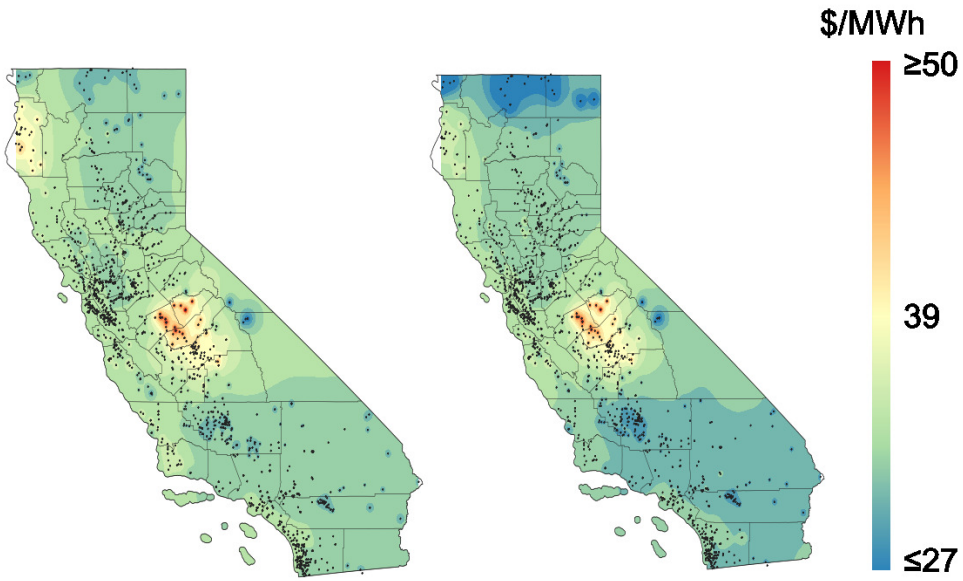


Figure 3.3: Temporal average price (at different spatial locations) for CAISO in the DAM (left) and RTM (right).

We conducted a basic statistical analysis to compute space-time price averages and standard deviations for the CAISO data set for the year of 2015. This dataset is open-access and was collected from CAISO Open Access Same-time Information System (OASIS) (AM and EC, 2014). The dataset includes complete electricity price profiles for the year at 2,234 different network locations. The data set contains over 19,569,840 price points for the DAM (one-hour time resolution) and 234,838,080 price points for the RTM (5-minute time resolution). We use this data to construct a space-time covariance matrix Σ form both the DAM and RTM.

The results are visualized in Figures 3.3, 3.4, and 3.5. Figure 3.3 illustrates that the time-average price for both markets is in the range of 27-50 USD/MWh. The space-time average RTM price is 32.71 USD/MWh, which is 2.62% lower than the corresponding average DAM price of 33.59 USD/MWh. The differences illustrate that there is a premium

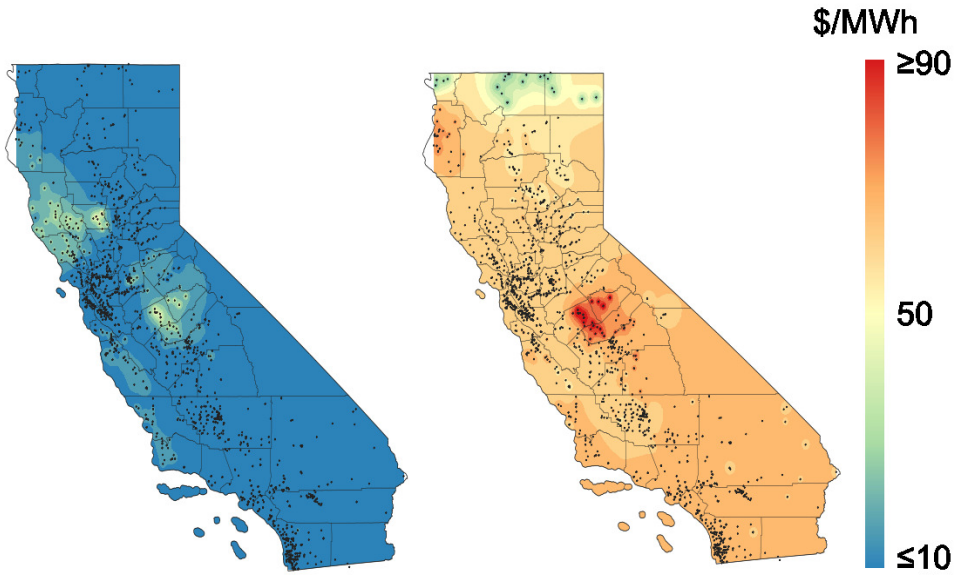


Figure 3.4: Temporal average standard deviation (at different spatial locations) for the DAM (left) and RTM (right).

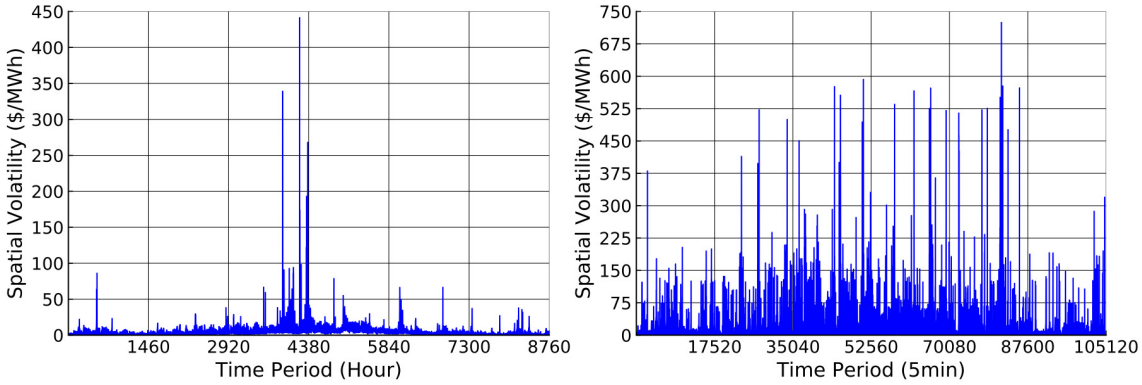


Figure 3.5: Spatial average standard deviation (at different temporal locations) for the DAM (left) and RTM (right).

in the DAM. Spatial patterns for both markets are quite similar, indicating that prices are dictated by the network topology. Figure 3.4 demonstrates temporal price volatility (standard deviation) at all locations. The temporal volatility in the DAM is consistently under 10 USD/MWh in most locations while the volatility in the RTM is in the range of 60-70 USD/MWh and reaches levels of 90 USD/MWh in some locations. The spatial average of the temporal volatilities was found to be 62.41 USD/MWh for the RTM, almost

four times larger than in the DAM, which was only 12.93 USD/MWh. These results clearly indicate that RTM possesses greater temporal volatility. Figure 3.5 presents spatial volatility through time. We see that the DAM shows low spatial volatility (except in a few instances in the summer months) while the RTM shows more frequent spikes in spatial volatility. Based on this analysis we conclude that the RTM is more volatile than the DAM in both time and space. We also found that the temporal average of the spatial volatility was found to be 8.85 USD/MWh for the RTM and 5.60 USD/MWh for the DAM. We can thus see that, on average, spatial volatility is less significant than temporal volatility (which are 12.93 USD/MWh for DAM and 62.42 USD/MWh for the RTM).

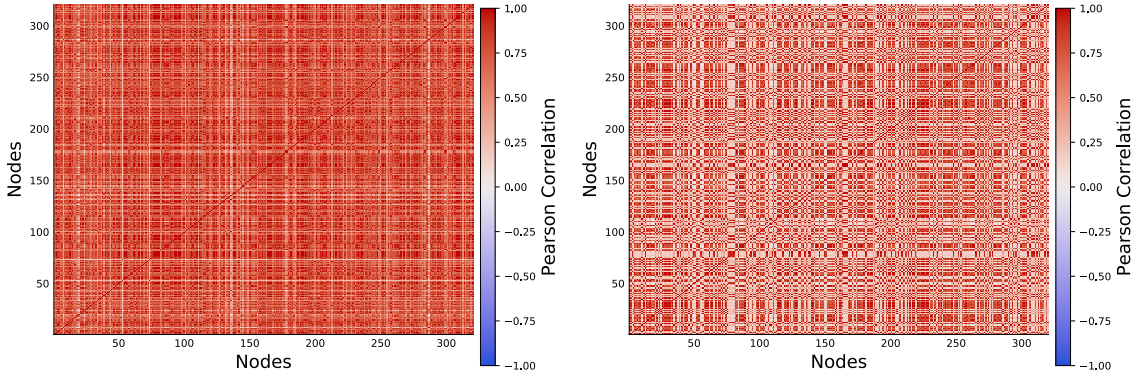


Figure 3.6: Pearson Correlation Matrix for the DAM (left) and RTM (right).

We also computed the spatial correlation matrix based on the covariance matrix and it is visualized in Figure 3.6. Our results show that, in the DAM, the average correlation is 0.67, that 99% of the total number of locations are positively correlated, and that the minimum correlation is -0.22. In the RTM, the average correlation is 0.82 and the minimum correlation is 0.00083. We conclude that a strong positive correlation exists in both electricity markets (prices at different locations tend to move in the same direction). This indicates that there is tight physical network coupling. As we will see next, strong positive correlation indicates that it is *impossible* to eliminate investment risk by simply investing in either generation or loads (a combination of both is needed). This would not be the case if we had a strong negative correlation in the market.

3.4.2 *Eigenvalue Analysis of Space-Time Covariance Matrix*

Solving the basic placement formulation is equivalent to solving an eigenvalue problem. In Table 3.3 and Figure 3.7, we summarize the eigenvalue spectrum in ascending order for both the DAM and RTM price covariance matrices (recall that the eigenvalues are the variances of the profit). Recall that both the DAM and RTM matrices have a total of 2,234 eigenvalues. The first 1,180 eigenvalues of the DAM price covariance are close to zero. For the RTM, the first 1,454 eigenvalues are close to zero (below a threshold value of $O(10^{-2})$). This indicates that many eigenvectors (allocations) give zero variance, meaning that many combinations of asset locations (given by the corresponding eigenvectors) can eliminate profit variance. An optimal strategy to eliminate risk is to place combinations of loads and generators at neighboring nodes (those with similar temporal price profiles). This can be visualized in Figure 3.8, where we show the optimal placement of assets (the eigenvectors) corresponding to the minimum eigenvalues. As can be seen, allocations of generation and load always appear in pairs next to each other and are of equal magnitude.

Table 3.3: Eigenvalues for DAM and RTM covariance matrices.

Eigenvalue	DAM	RTM
λ_1	-4.25×10^{-12}	-4.59×10^{-11}
λ_{10}	-2.09×10^{-14}	-1.90×10^{-12}
λ_{100}	-2.86×10^{-16}	-5.74×10^{-16}
λ_{500}	2.91×10^{-18}	5.80×10^{-17}
λ_{1000}	5.78×10^{-4}	4.68×10^{-5}
λ_{1500}	0.24	0.016
λ_{2000}	5.87	9.20
λ_{2100}	18.35	74.15
λ_{2200}	300.62	2806.04
λ_{2234}	2.54×10^{5}	7.31×10^{6}

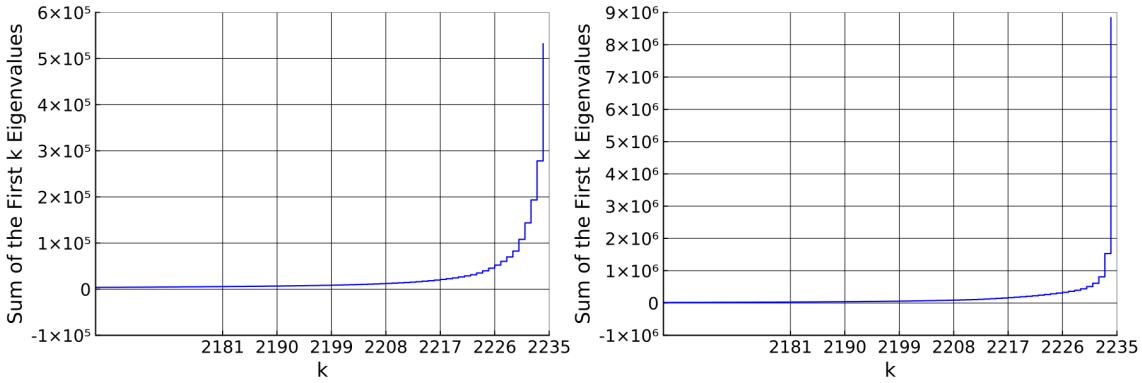


Figure 3.7: Cumulative eigenvalue spectrum for the DAM (left) and the RTM (right) covariances.

The largest eigenvalue (the maximum possible profit variance) is $O(10^5)$ for the DAM and $O(10^6)$ for the RTM, indicating that there is more volatility in the RTM (reinforcing the observations made with basic statistical analysis). In Figure 3.9, we present the optimal allocations corresponding to the maximum variance. The strategy here is to place the same asset type (in this case power generation) at all nodes. The maps also reveal areas the are strongly positively correlated (so the strategy to maximize variance is to allocate more generation at such locations). Obviously, this strategy is not optimal from an investment standpoint but highlights some interesting properties of the behavior of electricity prices.

3.4.3 *Risk vs. Mean Profit Trade-off for the DAM*

We used the placement formulation to analyze trade-offs between risk and expected profit. In Table 3.4 and Figure 3.10 we present the optimal trade-off solutions (Pareto optimal solutions) for the DAM. The Pareto solutions were identified using an ϵ-constrained approach. From these results we can make a number of interesting observations. First, it is clear that to maximize expected profit it is optimal to centralize facilities (these facilities are simply installed at locations with large mean price). In this case, obviously, the type of asset to install is generation and the expected profit is 52.76 \$/MWh. This strategy, however, results in a large risk (an MD value of 24.03 \$/MWh). We can also see that the

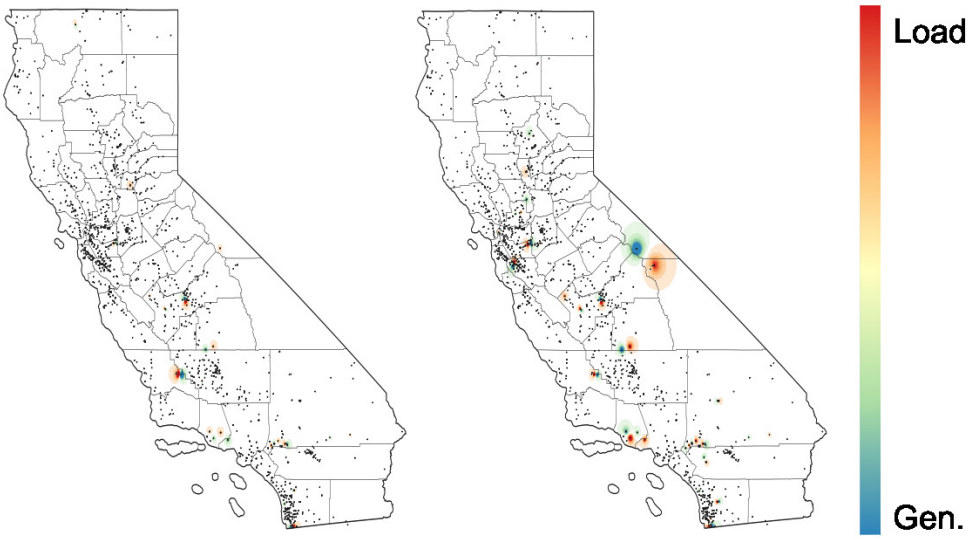


Figure 3.8: Optimal placement leading to zero risk for the DAM (left) and RTM (right).

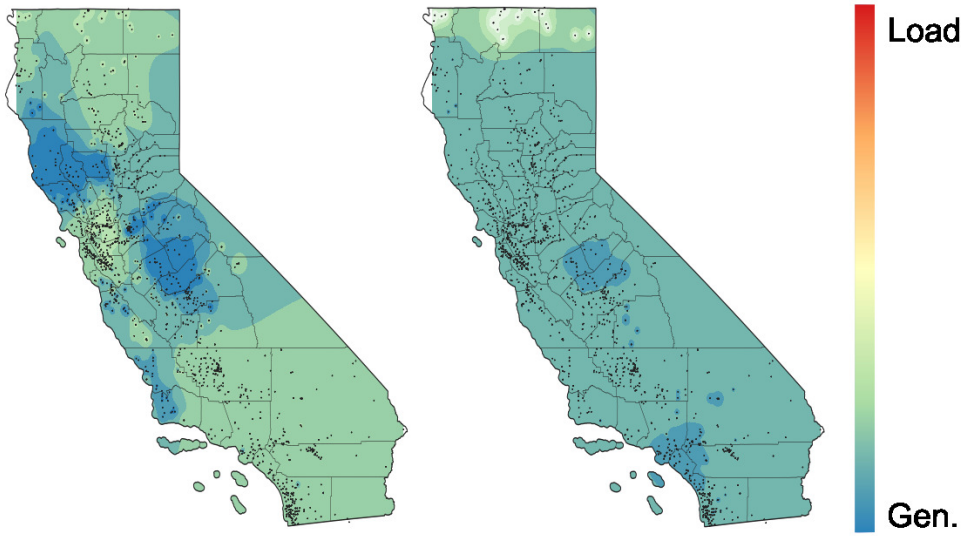


Figure 3.9: Optimal placement leading to maximum risk for the DAM (left) and RTM (right).

mean deviation is significant, representing half of the expected profit, which is due to the high temporal volatility of the prices. The trade-off trends also indicate that installation of a larger number of smaller power generators (diversifying generation among multiple locations) can substantially decrease risk. For instance, by increasing the number of generators to five, we see that the risk is decreased by 50% and this only decreases the

expected profit by 15%. This illustrates that there is a strong nonlinear trade-off between expected profit and risk. The mean absolute deviation gives linear penalties for large and small deviations (compared to the standard deviation, which attributes quadratic penalties). In Table 3.4 we present the standard deviation values for each placement problem solved. Note that the trend of standard deviation agrees with that of the mean absolute deviation. Therefore, choosing the mean absolute deviation as the risk measure for the optimal placement problem is consistent. In other words, one can recover elements of the Pareto frontier corresponding to the standard deviation by using the mean absolute deviation (there is a one-to-one corresponding between the risk measures).

From Table 3.4 and Figure 3.10 we see that further reductions in risk require the installation of both generation and loads. In particular, elimination of risk cannot be achieved through the use of either just generation or just loads (due to the positive correlation of prices). In the hypothetical case in which market prices were negatively correlated, installing the same asset type would be sufficient to fully mitigate risk. Consequently, the limiting value of risk for single asset type is an indicator of the degree of positive correlation in the market. Figure 3.11 shows optimal placement locations for low-risk and high-risk cases. We see that high-risk is achieved by placing only generation assets while low-risk is achieved by diversifying loads and generation.

An interesting trade-off point that we see in Table 3.4 is that in which we obtain a risk of MD = 5.06 USD/MWh and expected profit of μ_φ = 31.85 USD/MWh (this is the solution for minimum possible risk achieved with only generation assets). In this solution, seven generation locations achieve a mean absolute deviation of 5.06 MWh and an expected profit of 31.85 MW (78.94% of the risk is reduced while 39.63% of the profit is sacrificed). In Table 3.5 we show the power allocation to each of the seven locations. We see that two locations share 90% of the total generation (these seek to maximize expected profit) while 10% of the generation is split in small generators (these seek to minimize risk). From Table 3.4, we see that the use of just two generators incurs a large risk. Consequently, investing in smaller generators is key to mitigate risk. From these results

Table 3.4: Risk vs. expected profit trade-off for DAM.

Risk (USD/MWh)	Std. dev. (USD/MWh)	Expected Profit (USD/MWh)	# of Loads	# of Generators
24.03	46.66	52.76	0	2
18.15	33.78	49.15	0	3
12.19	19.52	44.22	0	5
8.31	11.64	39.24	0	12
5.75	7.85	34.16	0	8
5.06	7.00	31.85	0	7
4.25	5.53	28.15	3	9
3.15	4.08	22.13	13	9
1.33	1.75	10.63	29	12
0.16	0.23	1.92	142	112
0.055	0.011	0.80	287	254

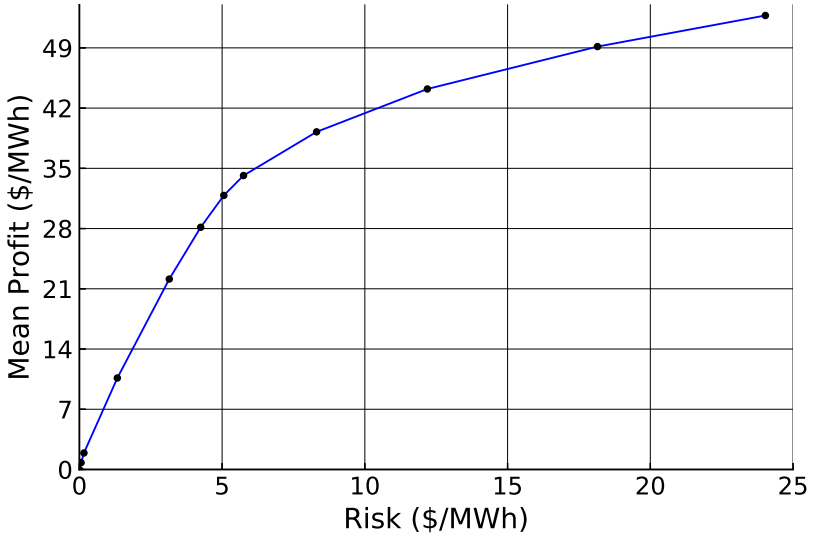


Figure 3.10: Risk vs. expected profit trade-off for the DAM.

we also conclude that further diversification of generation does not provide significant benefits in risk mitigation.

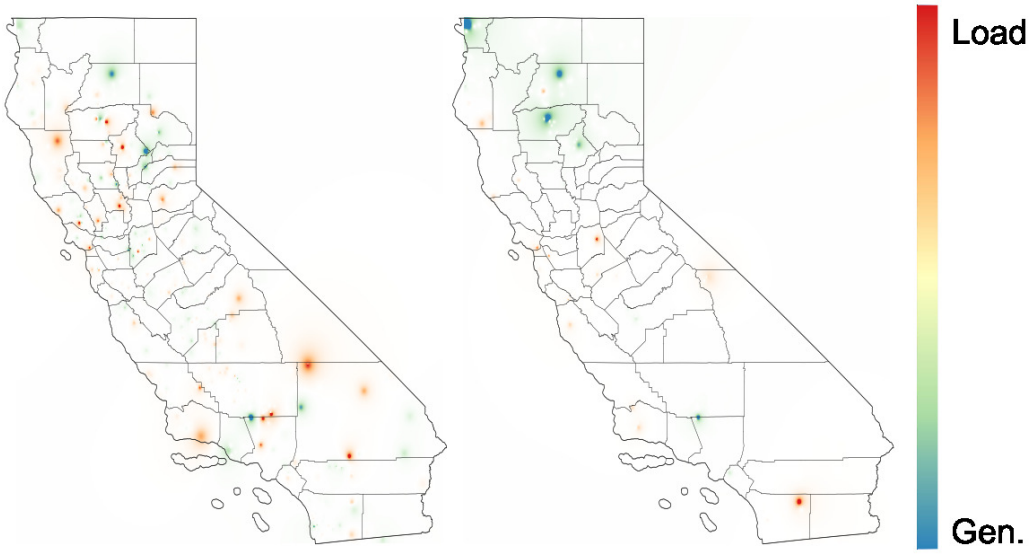


Figure 3.11: Optimal placement for low risk (left) and high risk (right) in the DAM.

Table 3.5: Optimal allocation for case with MD = 5.06 USD/MWh and $-\mu_\varphi$ = 31.85 USD/MWh in the DAM.

Location	w_i (MWh)
NEORBLF_7_B1	0.60
JBBLACK1_7_B1	0.31
DELNORTE_LNODED50	0.037
HMBUNIT2_7_GN010	0.023
HMBLTBY_6_N003	0.017
TOPAZC1_7_N021	0.010
BAFCOG12_7_B1	0.00090

3.4.4 *Risk vs. Expected Profit Trade-off for the RTM*

Trade-off analysis for the RTM was performed by using price data with a time resolution of 20 minutes. The reason is that the placement problem is intractable at higher resolutions. The Pareto analysis results are summarized in Table 3.6 and Figure 3.12. Here, we report standard deviation values in order to highlight how the mean and stan-

dard deviation follow the same trend. The results for RTM have similar trends to those found in the DAM. In contrast with the DAM results, however, the risk for RTM is higher (which is consistent with the results obtained using the eigenvalue analysis). Compared to the DAM, more diversification of generation is needed to decrease the risk by the same amount (due to the higher volatility in RTM). We also observe that a combination of loads and generators is needed to fully eliminate risk and that the expected profit obtained with the RTM and DAM are similar.

Figure 3.13 shows high-risk and low-risk allocations. High-risk allocations with large expected profit favor centralization of assets while low-risk ones favor decentralization of assets. Moreover, this indicates that assets capable of providing simultaneous provision of generation and load (e.g., microgrids or batteries) can be used to mitigate risk. Our analysis also indicates that electricity markets provide significant incentives to modularize power-intensive assets (e.g., manufacturing facilities and data centers). For instance, decentralization of ammonia systems can help mitigate risk associated with the high consumption of electricity in refrigeration systems.

The risk estimated with the 20-minute formulation underestimates that of the 5-minute counterpart. We can see, however, that the 20-minute resolution data already reveals that much higher risk is observed in RTM relative to DAM. This observation is also confirmed using the eigenvalue analysis (which was performed using the 5-minute resolution data). Moreover, we expect similar trade-off trends by using higher time resolutions.

3.4.5 *Computational Considerations*

The unconstrained placement problem is an eigenvalue problem that can be readily solved for both the DAM and the RTM data (even at 5 minute resolutions). The constrained placement formulation (3.3.15), on the other hand, is a large-scale mixed-integer linear program. The RTM problem (with 20-min resolution) contains 85,544 constraints

Table 3.6: Risk vs. expected profit trade-off for the RTM.

Risk (USD/MWh)	Std. dev. (USD/MWh)	Expected Profit (USD/MWh)	# of Loads	# of Generators
36.85	94.55	51.59	0	1
29	79.96	47.03	0	3
24	68.47	43.75	0	4
19	57.79	40.22	0	4
16	54.11	37.70	0	6
13	44.04	33.98	0	13
10	31.08	29.46	0	22
8	20.86	25.43	0	21
6	14.53	19.71	6	18
4	9.87	13.44	21	22
2	4.98	6.88	45	47
1	2.55	3.53	71	62
0.6	1.56	2.17	87	84
0.4	1.06	1.49	103	99
0.2	0.54	0.76	130	157

and 61,496 variables (4,468 binary) while the DAM problem contains 32,985 constraints, 21,989 continuous variable (4,468 binary). The problems were solved using `Gurobi` with default relative MIP gap of 0.01% and solution times range from 1.5 hours to 5 hours (on a standard personal computer). The long times are due to significant symmetries in the problem (i.e., many allocation combinations achieve the same optimal objective). This degeneracy was revealed by the eigenvalue analysis (which indicates that the price covariance has a large number of zero eigenvalues). The RTM problem is intractable with time resolutions below 20 minutes. We are currently investigating strategies to decompose the placement problem in order to be able to scale to higher time resolutions. In particular,

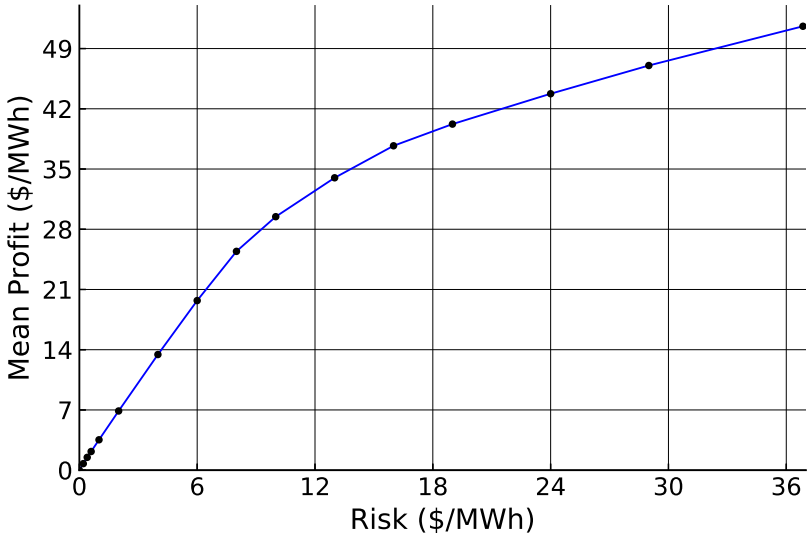


Figure 3.12: Risk vs. expected profit trade-off for the RTM.

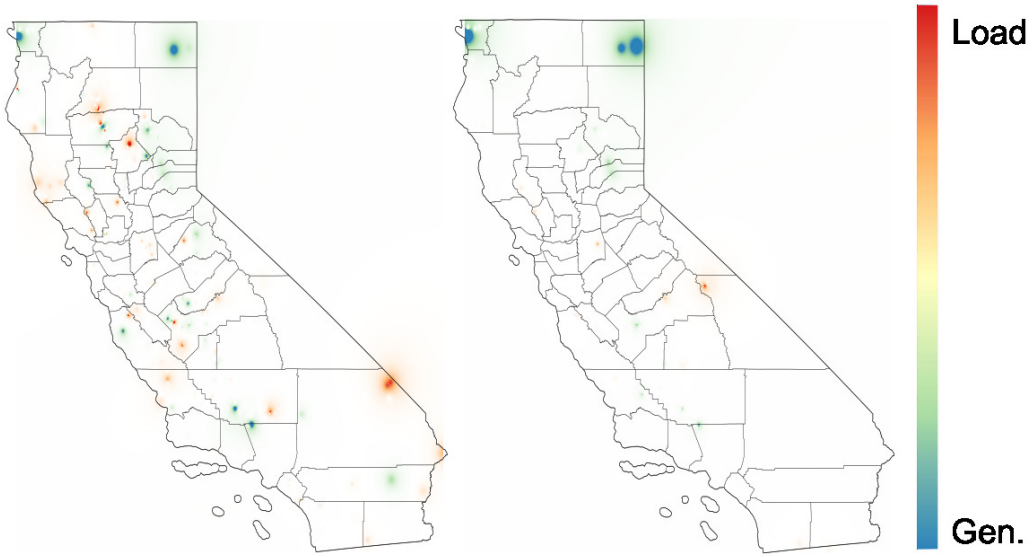


Figure 3.13: Optimal placement for low risk (left) and high risk (right) in the RTM.

this problem has the interesting property that it only has a single coupling constraint. Consequently, one can develop specialized Lagrangian decomposition (Fisher, 2004; Held and Karp, 1971) schemes that achieve high parallel execution efficiency.

3.5 Conclusions and Future Work

This work examines economic incentives created by space-time dynamics of day-ahead and real-time electricity markets. We first formulate the electricity market clearing model and explore the effect of modularization on social welfare and electricity prices. We then develop an optimal technology placement formulation that seeks to identify optimal strategies to maximize expected profit and minimize risk. We have shown that a pure risk minimization formulation can be cast as an eigenvalue problem. We also develop more sophisticated formulations that capture different technology asset types (e.g., generation or loads) and risk measures using mixed-integer programming techniques. Our analysis for the CAISO market reveals that significantly more temporal (as opposed to spatial) volatility is observed in both DAM and RTM markets (the RTM also has more volatility in general). Our analysis also reveals that both markets exhibit positive spatial correlation in prices, indicating that it is *impossible* to fully eliminate risk by using only either generators or loads. Consequently, decentralizing technologies of the same type has significant but limited impacts on risk mitigation. Full risk mitigation can only be achieved by combinations of generation and load assets (which can be achieved with microgrids, prosumers, or batteries). Our analysis also indicates that electricity markets provide significant incentives to modularize power-intensive technologies (e.g., manufacturing and data centers). This is of particular relevance due to recent interest in the deployment of small-scale modular technologies.

Our analysis is retroactive in nature (uses historical data) and thus lacks predictive capabilities. Enabling predictability requires us to develop advanced forecasting methods that capture simultaneous spatial and temporal correlations. Specifically, we are interested in investigating recently-developed dynamic principal component analysis and dynamic mode decomposition techniques to conduct space-time analysis and forecasting of market data (Vanhatalo et al., 2017; Dong and Qin, 2018). Such techniques exploit space-time correlations to identify dominant modes in the data.

As time points in RTM exceed 100 thousands, current state-of-art solvers could not resolve the placement problem under reasonable CPU time. Methods to shrink the problem size such as grouping the time points and reducing the number of redundant locations could be explored. Also, our model currently assumes that electricity price at each time point is independent, and therefore no temporal correlation of the electricity prices are considered. It is necessary to extend the placement formulations to analyze effects of temporal flexibility. In particular, our current formulation assumes that technologies provide a fixed capacity all the time, while new technologies can provide ramping capacity to account for uncertainty due to demand and renewable forecasting errors. Besides this, electricity storage facilities could potentially help to reduce risk in the volatile market. Therefore, we are interested in developing advanced formulations that can capture technologies that can shift load/generation in time. These formulations are intractable with off-the-shelf tools (due to a dramatic increase in the number of decision variables) and we will thus investigate decomposition algorithms for their solution. Specifically, such formulations reach tens to hundreds of millions of variables but note that the placement formulation exhibits sparse coupling (the total installed capacity constraint). As a result, one can envision using Lagrangian dual decomposition techniques to tackle this problem. We are also interested in including nonlinear effects of economy of scales in the placement, which can be done by using piece-wise linear approximations.

4

BENEFITS OF MODULAR DESIGN - TEMPORAL FLEXIBILITY

This chapter talks about the temporal flexibility brought by modular technologies, and is published in Shao et al. (2021) titled "Mitigating Investment Risk Using Modular Technologies".

4.1 Introduction

Decentralized power generation and storage systems are becoming increasingly attractive as climate change and adoption of renewable power disrupts markets and space-time demand patterns (Heuberger et al., 2017; Liu et al., 2018; Shao and Zavala, 2019). Modular technologies can be easily transported to different geographical locations to exploit changing market patterns and to enable the recovery of resources that are highly distributed and potentially short-lived (Allman and Zhang, 2020; Chen and Grossmann, 2019; Davis, 2016). We can interpret this ability as a form of *spatial-shifting* flexibility. This decentralized approach contrasts with the more traditional monolithic approach in which a large processing system is installed at a fixed location over its entire lifetime (Zhao et al., 2018). This centralized approach involves investments that can reach billions of US dollars and face significant risk due to changing markets and climate, shortages of resources at a spe-

cific location (e.g., water), and changes in the policy landscape (e.g., carbon emissions). As such, large central systems can face significant economic fallouts that investors might not be willing to tolerate. For instance, large ammonia production systems in the US have shut down due to low-cost supply from China, and large coal power plants are shutting down due to decreasing costs of renewable power and policy changes. Moreover, the mass deployment of small modular units facilitates experimentation, learning, and sharing of best practices that can reduce operational costs (compared to large facilities, in which experimentation is more difficult). On the downside, the flexibility provided by small modular systems often comes at the expense of increased investment and operational costs (Rajagopalan, 1993). Specifically, economies of scale benefit large systems due to the favorable scaling of throughput with equipment size (Peters et al., 1968). Due to complex trade-offs between costs and flexibility, industrial systems will likely evolve into a mixed state in which certain processing tasks are performed in small modular systems while others are performed in large centralized systems. Identifying optimal investment strategies in such settings is complicated due to complex product interdependencies and uncertainties.

A key observation driving this work is that modular systems provide logistical flexibility in investment size and timing that can be strategically exploited to mitigate risk. Specifically, expansion of production capacity in modular systems can proceed sequentially, which provides a mechanism to hedge against risk (we can interpret this as *temporal-shifting* flexibility). To give an example, the deployment of new power generators and transmission lines is subject to significant short-term and long-term uncertainties. Specifically, short-term fluctuations in demand and wind/solar supply can affect an optimal generation mix, and changes in fuel prices and policy can render entire technologies uneconomical (Liu et al., 2018). Therefore, the progressive expansion of capacity using both large and small processing systems can help make and correct decisions and to better balance cost and risk.

In this work, we investigate investment flexibility provided by modular technologies;

to do so, we propose a multi-product capacity expansion (CE) problem that exploits the availability of technologies of different types and sizes to mitigate risk. Variants of the CE problem have been studied in different applications such as power generation, semiconductor manufacturing, railroad networks, and waste-to-energy systems (Cardin and Hu, 2015; Sun and Schonfeld, 2015; SHIINA et al., 2018; Geng et al., 2009). A cost-minimization CE problem that considers a single-product deterministic setting with installation decisions of a fixed-capacity facility was formulated in (Luss, 1979). This formulation was extended to incorporate facilities with multiple capacities in (Luss, 1983, 1986). Uncertainty in demand for a single-product cost-minimization CE problem was addressed by using a stochastic programming (SP) model in (F.H. Murphy and Soyster, 1982; DapkLus and Bowe, 1984; Shiina and Birge, 2003). A stochastic CE formulation for planning investments in electricity generation, storage, and transmission investments over a long planning horizon was proposed in (Liu et al., 2018). These CE problem formulations use expected cost as an investment metric and thus do not control investment risk. Recently, a CE problem formulation that trades-off expected cost and risk was proposed in(Zhao et al., 2019). Here, the conditional value-at-risk (CVaR) was used as a risk metric that is minimized at each stage. All the aforementioned formulations consider facilities that produce a single product; in a chemical process, however, multi-product dependencies need to be captured. Specifically, a chemical manufacturing facility might involve processes that produce intermediate or final products and demands for such products might face different levels of uncertainty. Making investment decisions in a multi-product setting is a non-trivial problem. Capturing risk in time-dependent decision-making settings (such as CE) is also an active topic of research. For instance, time-consistency of per-stage risk minimization is an issue of concern. In the context of CE, time consistency indicates that, if an alternative A is riskier than alternative B at some time, then A should also be considered riskier than B at every prior time (Boda and Filar, 2006). Unfortunately, deriving SP formulations that achieve time-consistency is not straightforward. Moreover, per-stage risk minimization is not necessarily a decision-making strategy that

investors might follow; specifically, investors are typically concerned with assessing risk of cumulative metrics such as the net present value (NPV).

In this work, we propose a multi-product CE formulation to investigate flexibility brought by modularization for mitigating investment risk. Our framework is a multi-stage and multi-objective SP problem that captures demand product uncertainty and trade-offs between expected value and risk of the NPV. We provide case studies of different complexity to illustrate the developments. Our analysis reveals that the Pareto frontier of a flexible setting (allowing for deployment of units of various sizes) dominates the Pareto frontier of an inflexible setting (allowing only for deployment of large units). Our formulation also avoids difficulties associated with time-consistency issues of stage-wise risk-minimization formulations and we argue that is more compatible with more traditional investment strategies.

4.2 Problem Formulations

In this section, we present CE formulations of different complexity (single-product/multi-product and deterministic/stochastic) in order to highlight different aspects of the problem. We begin our discussion by posing a couple of illustrative examples; this will help us introduce some key concepts that are essential in developing more complex CE formulations.

4.2.1 *Problem Setting*

Consider the following deterministic CE setting: a decision-maker (investor) wants to progressively add capacity to a production system by installing technologies of different sizes (capacities). The resulting assembled system seeks to generate sufficient product to satisfy a time-dependent demand over a given planning horizon. At each planning stage, the investor decides how many technologies (and associated capacities) it should install; if a technology is added at one stage, this will generate a product to satisfy the

demand at the next stage (there is a deployment delay of one stage). Demand satisfaction generates revenue. We assume that an installed technology has to operate at full capacity; if the system production exceeds demand at a given time, the investor can decide to either store the excess product at a cost (and carry the product over to the next stage) or it can dispose of excess product at a cost. At the final stage, the system disposes of leftover excess product. The goal is to make an optimal CE plan over the horizon that maximizes NPV (accumulated cash flows over the horizon); in doing so, the investor is constrained by the capacities of the technologies available. For simplicity, in this example, we assume that NPV is simply determined by the excess product (waste) at the end of the planning horizon and that there is no interest rate.

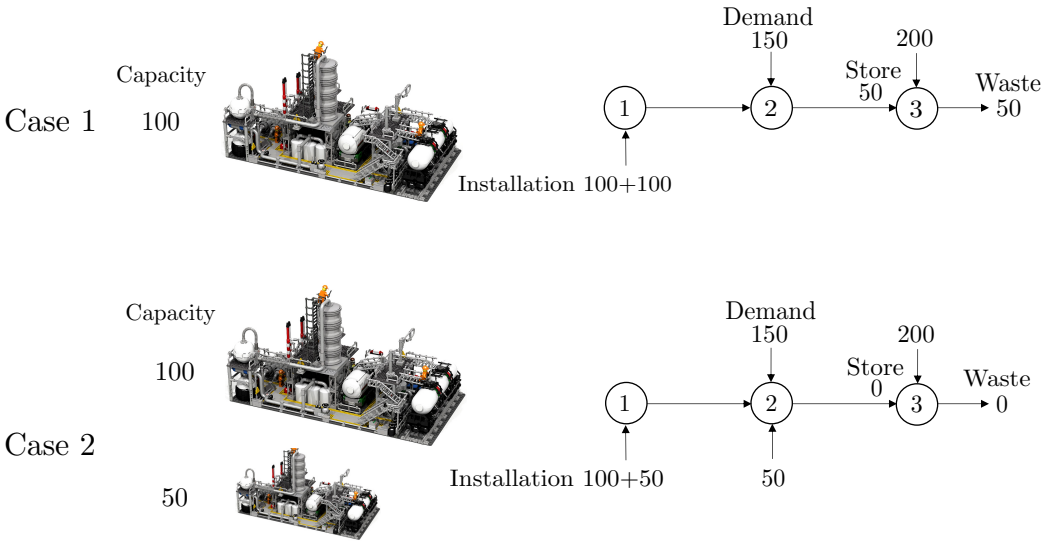


Figure 4.1: Illustrative example of the single-product deterministic capacity expansion setting.

We illustrate this decision-making setting in Figure 4.1; here, we would like to make decisions on how much capacity to install at Stage 1 and Stage 2 to minimize waste at Stage 3. In Case 1, only large technologies are available (with a capacity of 100 units); to satisfy future demands, it is decided to install 2 units of this large technology at Stage 1. Since the demand at Stage 2 is 150, it is required to shift excess production to Stage 3. Moreover, since the demand at Stage 3 is 200, it is necessary to dispose of 50 units of

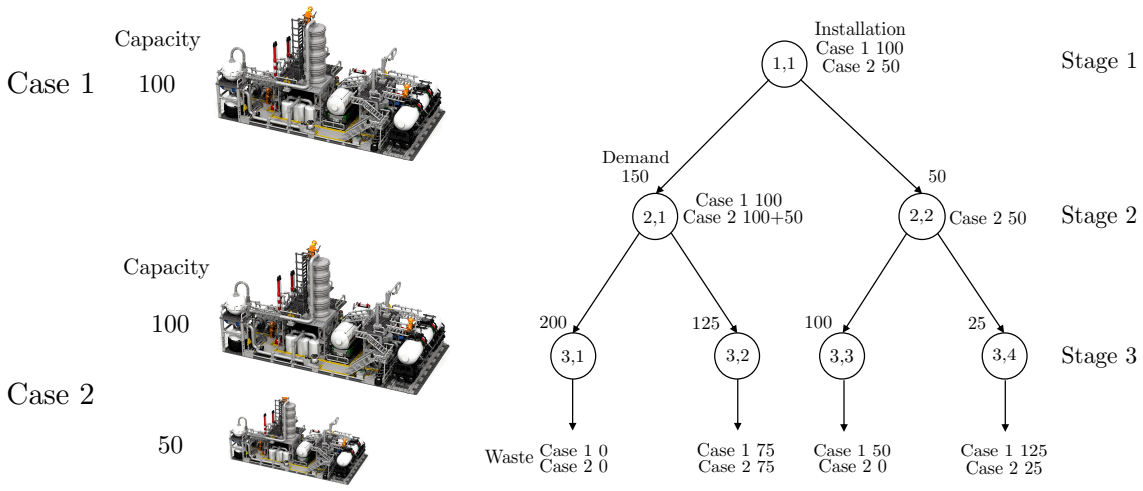


Figure 4.2: Illustrative example of the single-product stochastic capacity expansion setting.

excess product. In Case 2, a large technology and a small technology are available; this opens the possibility of an investment strategy in which we install a large unit (size 100) and a small unit (size 50) at Stage 1 and add a small unit at Stage 2. This strategy prevents wasting material at Stage 3 and highlights the flexibility provided by the availability of small technologies. Note that, in this setting, the demands are time-dependent but are assumed to be known at the moment of decision (deterministic setting).

The CE setting can be extended to account for uncertainty in the demands (stochastic setting). Here, demand uncertainty is represented in the form of possible scenarios. An illustrative example of this setting is shown in Figure 4.2. We would like to make installation decisions at each stage and scenario (here, we consider two possible scenarios per stage). Stages and scenarios are represented as a decision tree and each node is associated with a different demand scenario. Installation decisions are shown next to the node and wasted amounts are shown exiting the nodes at the last stage. In Case 1 (only large technologies are available), we decide to install a large technology in Stage 1; in Stage 2, we can decide to install a large technology in scenario 1 (high demand) or no technology in Scenario 2 (low demand). This investment strategy results in four scenarios of waste product in Stage 3 (10,75,50,125). Assuming that these scenarios have equal probability

(1/4), the expected value of the waste is 65 and the standard deviation (typical measure of risk) is 48. In Case 2 (large and small technologies available), we install a small unit in Stage 1; in Scenario 1 in Stage 2 we install a large technology (to satisfy the large demand) and in Scenario 2 we install another small technology (to satisfy the small demand). This investment strategy results in four scenarios of waste excess product in Stage 3 (0,75,0,25). This gives a mean waste of 25 and a risk of 35. We can thus see that adding the possibility of installing small units reduces expected waste and risk.

Risk can be measured in different ways; in the previous setting, we computed the risk at Stage 3 (last stage) but we could have also computed the risk at Stage 2 and we could have added this to the risk of Stage 3 (add risks for all stages) to determine the best strategy. This highlights issues that one may encounter when measuring risk in a multi-stage decision-making setting. Specifically, risk can vary over time and one might or might not be interested in shaping risk over time. This is similar in spirit to how investors think about cash flows; typically, investors are not necessarily interested in the temporal behavior of cash flows but want to aggregate cash flows in a single metric (e.g., NPV). Following this reasoning, in this work, we will compute NPV for every branch in the tree and compute the associated risk.

The CE problem can be further extended to a multi-product setting in which a system can produce multiple intermediate or final products. Intermediate products generate interdependencies between possible technologies (i.e., technology can take intermediate products obtained from another technology as raw materials). Multi-product dependencies make the problem significantly more complicated and we will see that, in such a setting, investment flexibility provided by small units becomes particularly relevant. We now proceed to formulate single-product deterministic and stochastic CE problems and we then proceed to extend this to a multi-product setting.

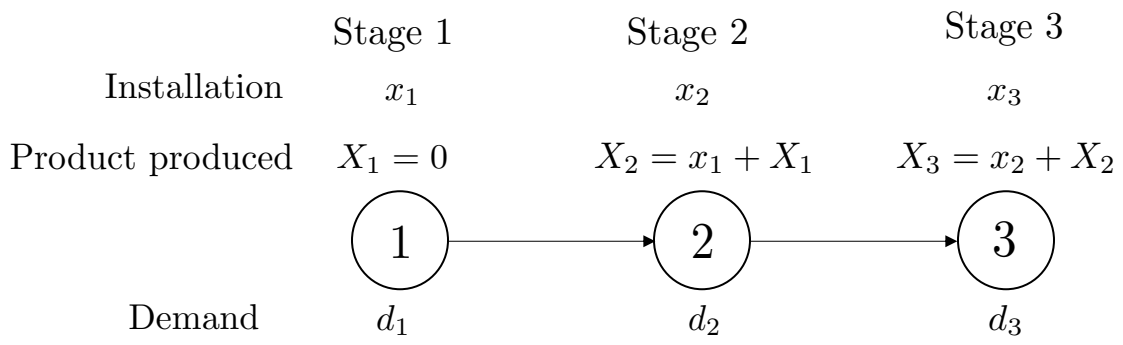


Figure 4.3: Tree representation of planning horizon in deterministic case

4.2.2 *Single-Product, Deterministic Setting*

Consider the decision-making setting shown in Figure 4.3. We consider a planning horizon comprising a set of stages $\mathcal{T} = \{1, 2, \ldots, T\}$ with cardinality $|\mathcal{T}| := T$. The time-dependent product demand is given by $d_t, t \in \mathcal{T}$. Investment decisions are made at stages $t \in \{1, 2, \ldots, T-1\}$ and we thus define the decision stages $\mathcal{D} = \{1, 2, \ldots, T-1\}$ with cardinality $|\mathcal{D}| := T - 1$. In a deterministic setting, the planning horizon is a linear graph (a tree) in which each node represents a stage. As such, for each node t, we define a parent node $a_t \in \mathcal{T}$ (in this case we have $a_t = t - 1$). The root node $t = 1$ does not have a parent node and thus $a_1 = \emptyset$.

The investor has a list of of possible technology choices that can be installed at each stage. Each choice has a different capacity and associated installation cost (which capture economies of scale). We define the set of capacities as $\mathcal{B} = \{B_1, B_2, \ldots, B_N\} \in \mathbb{Z}_+^N$ and the set of associated costs as $\mathcal{C} = \{C_1, C_2, \ldots, C_N\} \in \mathbb{R}_+^N$, both with same cardinality $|\mathcal{B}| = |\mathcal{C}| := N$. For convenience, we also define a set of choice indexes $\mathcal{F} = \{1, 2, \ldots, N\}$. To capture economies of scale, it is typical to assume that costs follow a *2/3 rule* and thus:

$$\left(\frac{B_k}{B_{k'}}\right) = \left(\frac{C_k}{C_{k'}}\right)^{\frac{3}{2}}, \quad k, k' \in \mathcal{F} \tag{4.2.1}$$

where $B_k, B_{k'} \in \mathcal{B}$ are the k^{th} and k'^{th} capacity choices and $C_k, C_{k'} \in \mathcal{C}$ are the installation

costs.

Product storage comes at a cost $\rho_s \in \mathbb{R}+$ and we define a maximum storage capacity $\bar{s} \in \mathbb{Z}_+$. Disposal of excess product comes at a cost ρ_w. We define a variable $s_t \in \mathbb{Z}_+, t \in \mathcal{T}$ to capture the amount of storage at stage t. We set $s_1 = 0$ and $s_n = 0$ (any excess product is regarded as waste at the final stage). We define the integer variable $w_t \in \mathbb{Z}_+, t \in \mathcal{T}$ to represent the waste generated at each stage. We assume $w_1 = 0$ (waste is generated at the end of each stage). The investor has a choice to deal with any excess product; either to dispose of the product or to store it (shift it to the next stage). To capture installation *delays*, we assume that capacity installed at stage t generate production, storage, disposal and sales of products at stage $t + 1$.

We define integer variables $u_{t,k} \in \mathbb{Z}_+, t \in \mathcal{D}, k \in \mathcal{F}$; here, $u_{t,k}$ is the number of technologies of type $k \in \mathcal{B}$ installed at stage $t \in \mathcal{D}$. The total capacity installed at time t is thus:

$$x_t = \sum_{k \in \mathcal{F}} u_{t,k} B_k, \quad t \in \mathcal{D} \tag{4.2.2}$$

and the total installation cost at time t is:

$$y_t = \sum_{k \in \mathcal{F}} u_{t,k} C_k, \quad t \in \mathcal{D} \tag{4.2.3}$$

We define variable $X_t, t \in \mathcal{T}$ to represent the total amount of product generated at stage t; X_t is the cumulative installed capacity up to stage t and follows the dynamic evolution:

$$X_t = X_{a_t} + x_{a_t}, \quad t \in \mathcal{T} \tag{4.2.4}$$

We set the initial production as $X_1 = 0$ and recall that $a_t = t - 1$. In the proposed setting, we can install more than one technology at each time but we limit the total final installed capacity X_T by using the upper bound $\bar{x} \in \mathbb{Z}_+$.

We use the following constraint to ensure that demand is satisfies at each stage:

$$0 \leq s_{a_t} + X_t - s_t - w_t \leq d_t, \quad t \in \mathcal{T} \tag{4.2.5}$$

We define a production cost as $\rho_p \in \mathbb{R}_+$ and selling price as $\pi_p \in \mathbb{R}_+$. Under these definitions, the cost at stage t (denoted as $q_t \in \mathbb{R}$) can be expressed as:

$$q_t = y_t + \rho_p X_t + \rho_s s_t + \rho_w w_t, \quad t \in \mathcal{T}. \tag{4.2.6a}$$

and the revenue at stage t (denoted as $r_t \in \mathbb{R}$) is expressed as:

$$r_t = \pi_p(X_t + s_{a_t} - s_t - w_t), \quad t \in \mathcal{T}. \tag{4.2.7}$$

Note that $s_0, X_1, s_1, w_1 = 0$ and thus $r_1 = 0$.

We consider a CE formulation that maximizes the NPV of the project; to do so, we define an interest rate $\gamma \in [0, 1]$ that is used to discount any future cash flow and we define the discount factor $\beta_t = 1/(1+\gamma)^{t-1}$. We define the discounted profit (cash flow) achieved at stage t as v_t and the cumulative profit upto stage t as V_t. These quantities are computed as:

$$v_t = \beta_t \cdot (r_t - q_t), \quad t \in \mathcal{T} \tag{4.2.8a}$$

$$V_t = V_{a_t} + v_t, \quad t \in \mathcal{T}. \tag{4.2.8b}$$

With this, the NPV is given by $V_T = \sum_{t \in \mathcal{T}} v_t$.

In summary, the CE problem is a mixed-integer linear program (MILP) of the form:

$$\max_{u,s} \; V_T \tag{4.2.9a}$$

$$\text{s.t. } x_t = \sum_{k \in \mathcal{F}} u_{t,k} B_k, \;\; t \in \mathcal{D} \tag{4.2.9b}$$

$$y_t = \sum_{k \in \mathcal{F}} u_{t,k} C_k, \;\; t \in \mathcal{D} \tag{4.2.9c}$$

$$X_t = X_{a_t} + x_{a_t}, \;\; t \in \mathcal{T} \tag{4.2.9d}$$

$$q_t = y_t + \rho_p X_t + \rho_s s_t + \rho_w w_t, \;\; t \in \mathcal{T} \tag{4.2.9e}$$

$$r_t = \pi_p (X_t + s_{a_t} - s_t - w_t), \;\; t \in \mathcal{T} \tag{4.2.9f}$$

$$v_t = \beta_t \cdot (r_t - q_t), \;\; t \in \mathcal{T} \tag{4.2.9g}$$

$$V_t = V_{a_t} + v_t, \;\; t \in \mathcal{T} \tag{4.2.9h}$$

$$0 \leq s_{a_t} + X_t - s_t - w_t \leq d_t, \;\; t \in \mathcal{T} \tag{4.2.9i}$$

$$0 \leq s_t \leq \bar{s}, \;\; t \in \mathcal{D} \tag{4.2.9j}$$

$$X_T \leq \bar{x} \tag{4.2.9k}$$

$$u_{t,k} \in \mathbb{Z}_+, \;\; t \in \mathcal{D}, k \in \mathcal{F} \tag{4.2.9l}$$

$$s_t \in \mathbb{Z}_+, \;\; t \in \mathcal{T}. \tag{4.2.9m}$$

The NPV metric accumulates all the cash flows v_t, $t \in \mathcal{T}$ to the initial stage $t = 1$ and this accounts for time value of money. If we set $\gamma = 0$, we obtain $\beta_t = 1$ and the CE problem maximizes the cumulative cash flows over the planning horizon (the total profit). As we discuss next, the NPV is a convenient metric that allows us to summarize random cash flows that arise in settings that face uncertainty.

4.2.3 *Single-Product, Stochastic Setting*

We now extend the CE problem to a stochastic setting; this formulation allows us to explore trade-offs between expected profit and risk. The stochastic setting is illustrated

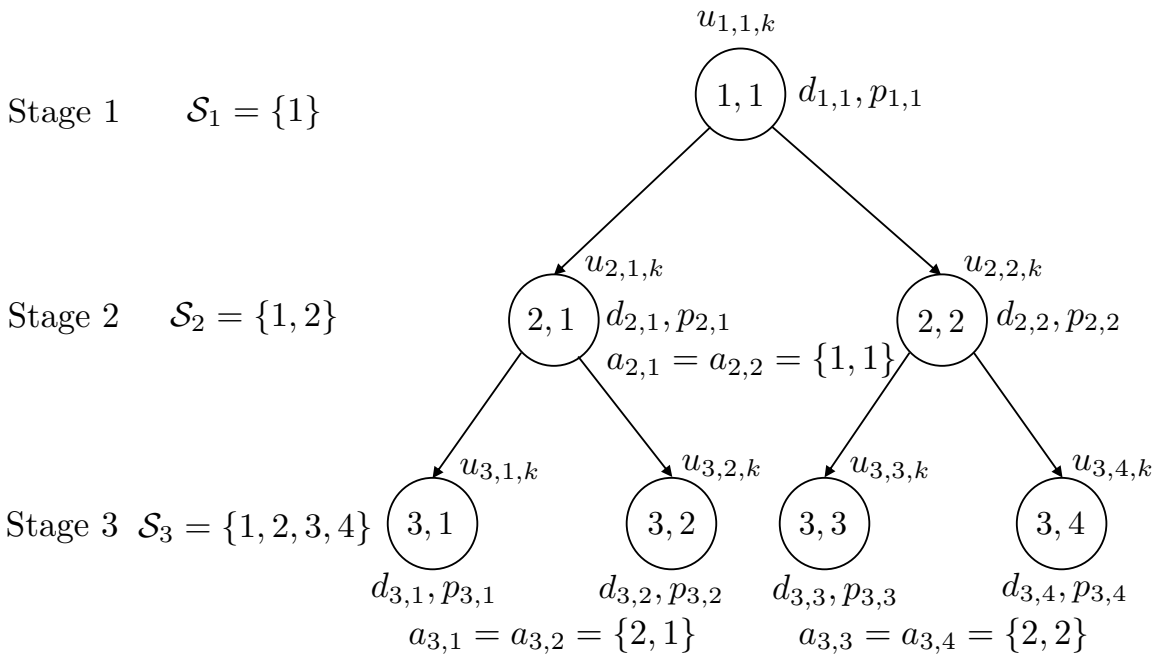


Figure 4.4: Tree representation of planning horizon in stochastic case

in Figure 4.4. We define the set of scenarios at each stage $t \in \mathcal{T}$ as $\mathcal{S}_t = \{1, 2, \dots S_t\}$ with cardinality $|\mathcal{S}_t| := S_t$. Each scenario is represented as a node in a tree; the number of levels in the tree is given by the number of stages. We define parent node $a_{t,j}, t \in \mathcal{T}, j \in \mathcal{S}_t$ as the parent stage and scenario that node $\{t, j\}$ emanates from. For example, if scenario $\{t, j_1'\}$ is generated from scenario j at stage $t-1$, then the parent node is $a_{t,j_1'} = \{t-1, j\}$ (see Figure 4.5). The scenario set of the root node is a singleton $\mathcal{S}_1 = \{1\}$ and the parent of the root node is empty and thus $a_{1,1} = \emptyset$.

The demand is a discrete random variable; the realization of this variable at time t and scenario j is denoted as $d_{t,j}, t \in \mathcal{T}, j \in \mathcal{S}_t$. The probability of realization $d_{t,j}$ is represented as $p_{t,j} \in [0, 1]$. For each stage $t \in \mathcal{T}$, these probabilities satisfy $\sum_{j \in \mathcal{S}_t} p_{t,j} = 1$. It is important to highlight that these are *joint* probabilities that capture the *history of events* leading to node $\{t, j\}$. In other words, joint probabilities are node probabilities (conditional probabilities are arc/edge probabilities and marginal probabilities ignore history). For example, in Figure 4.4, $p_{3,1}$ is the probability of node $\{3, 1\}$ corresponding

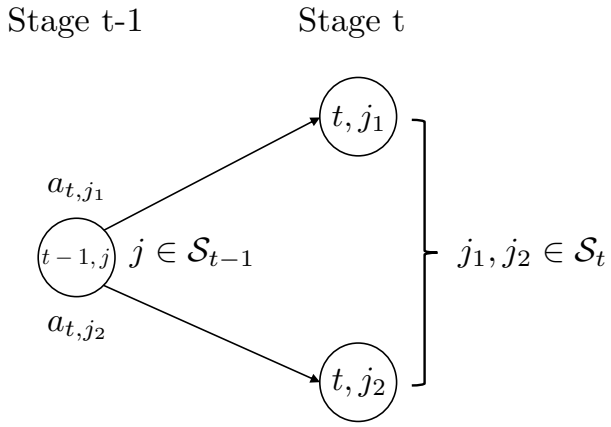


Figure 4.5: Schematic of parent-node notation. Here, node $\{t-1, j\}$ is the ancestor of $\{t-1, j_1'\}$ and $\{t-1, j\}$ and thus $a_{t,j_1} = a_{t,j_2} = \{t-1, j\}$.

to the demand event sequence $d_{1,1}, d_{2,1}$ and $d_{3,1}$ and thus:

$$p_{3,1} = \mathbb{P}(\{1,1\}, \{2,1\}, \{3,1\}) \tag{4.2.10a}$$

$$= \mathbb{P}(\{3,1\}|\{1,1\}, \{2,1\}) \cdot \mathbb{P}(\{1,1\}, \{2,1\}) \tag{4.2.10b}$$

$$= \mathbb{P}(\{3,1\}|\{1,1\}, \{2,1\}) \cdot p_{2,1} \tag{4.2.10c}$$

$$= \mathbb{P}(\{3,1\}|\{1,1\}, \{2,1\}) \cdot \mathbb{P}(\{2,1\}|\{1,1\}) \cdot p_{1,1}. \tag{4.2.10d}$$

where $\mathbb{P}(\{3,1\}|\{1,1\}, \{2,1\})$ is the conditional probability of event $\{3,1\}$ given that $\{1,1\}$ and $\{2,1\}$ have been realized. We thus have that probability $p_{t,j}$ carries information of its ancestor nodes (i.e., $p_{3,1}$ carries information of $p_{2,1}$ and $p_{1,1}$).

We define an integer variable $u_{t,j,k} \in \mathbb{Z}_+, t \in \mathcal{D}, j \in \mathcal{S}_t, k \in \mathcal{F}$ to represent the number of technologies of type $k \in \mathcal{B}$ installed at stage $t \in \mathcal{T}$ and at scenario $j \in \mathcal{S}_t$. The total capacity installed at time t and scenario j is:

$$x_{t,j} = \sum_{k \in \mathcal{F}} u_{t,j,k} B_k, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, \tag{4.2.11}$$

and total installation cost at time t and scenario j is:

$$y_{t,j} = \sum_{k \in \mathcal{F}} u_{t,j,k} C_k, \quad t \in \mathcal{D}, \quad j \in \mathcal{S}_t. \tag{4.2.12}$$

We use the integer variable $s_{t,j} \in \mathbb{Z}_+, t \in \mathcal{T}, j \in \mathcal{S}_t$ to denote the amount of storage at stage t in scenario j. Similar to the deterministic case, the storage for the first and last stage are assumed to be zero. We define the waste variable as $w_{t,j} \in \mathbb{Z}_+, t \in \mathcal{T}, j \in \mathcal{S}_t$ and we assume that the waste for the first stage zero.

We define a variable $X_{t,j}, t \in \mathcal{T}, j \in \mathcal{S}_t$ as the total amount of product produced at stage t and scenario j. Here, $X_{t,j}$ is interpreted as the cumulative installed capacity up to stage and scenario $a_{t,j}$, expressed in (4.2.13). The total installed capacity at stage one is $X_{1,j} = 0, j \in \mathcal{S}_1$.

$$X_{t,j} = X_{a_{t,j}} + x_{a_{t,j}}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.13}$$

With the above definitions, we can define the undiscounted cost at stage t for scenario j as $q_{t,j}$, and is given by:

$$q_{t,j} = y_{t,j} + \rho_p X_{t,j} + \rho_s s_{t,j} + \rho_w w_{t,j}, \quad t \in \mathcal{T} \setminus \{T\}, j \in \mathcal{S}_t. \tag{4.2.14a}$$

The undiscounted revenue at stage t scenario j is denoted as $r_{t,j}$ and can be expressed as

$$r_{t,j} = \pi_p (X_{t,j} + s_{a_{t,j}} - s_{t,j} - w_{t,j}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t. \tag{4.2.15a}$$

Our goal is to maximize the expected NPV and to minimize its risk. To model these quantities, we introduce variable $v_{t,j}, t \in \mathcal{T}, j \in \mathcal{S}_t$ that denotes the cash flow (profit) achieved in stage t and scenario j. We also define the cumulative variable $V_{t,j}, t \in \mathcal{T}, j \in \mathcal{S}_t$ to denote the cumulative profit up to stage t and scenario j. Using the notation proposed,

these quantities can be computed using a form that is analogous to the deterministic case:

$$v_{t,j} = \beta_t \cdot (r_{t,j} - q_{t,j}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.16a}$$

$$V_{t,j} = V_{a_{t,j}} + v_{t,j}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t. \tag{4.2.16b}$$

The NPV is the total accumulated cash flow and is given by $V_{T,j}$, $j \in \mathcal{S}_T$. We note that this is a random quantity and that each realization correspond to a branch of the scenario tree connecting the root node $\{1, 1\}$ to the final nodes $\{T, j\}$ with $j \in \mathcal{S}_T$. The NPV thus summarizes information of the entire project and captures probabilities of the different paths that the project can take. When the interest rate is zero, the NPV of a given path is the total profit of the project for such path.

The expected NPV is given by:

$$\mathcal{E} = \sum_{j \in \mathcal{S}_T} p_{T,j} V_{T,j}, \tag{4.2.17}$$

and its risk is measured by using the mean deviation:

$$\mathcal{R} = \sum_{j \in \mathcal{S}_T} p_{T,j} |V_{T,j} - \mathcal{E}|. \tag{4.2.18}$$

Alternative risk metrics can be used; here, we provide the mean deviation as this is a coherent risk measure that is easy to interpret.

The CE problem can be cast as the following stochastic, multistage, multi-objective optimization (SMMO) problem:

$$\max_{u,s} \ \{\mathcal{E}, -\mathcal{R}\} \tag{4.2.19a}$$

$$\text{s.t. } x_{t,j} = \sum_{k \in \mathcal{F}} u_{t,j,k} B_k, \quad t \in \mathcal{D}, j \in \mathcal{S}_t \tag{4.2.19b}$$

$$y_{t,j} = \sum_{k \in \mathcal{F}} u_{t,j,k} C_k, \quad t \in \mathcal{D}, \quad j \in \mathcal{S}_t \tag{4.2.19c}$$

$$X_{t,j} = X_{a_{t,j}} + x_{a_{t,j}}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.19d}$$

$$q_{t,j} = y_{t,j} + \rho_p X_{t,j} + \rho_s s_{t,j} + \rho_w w_{t,j}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.19e}$$

$$r_{t,j} = \pi_p (X_{t,j} + s_{a_{t,j}} - s_{t,j} - w_{t,j}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.19f}$$

$$v_{t,j} = \beta_t (r_{t,j} - q_{t,j}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.19g}$$

$$V_{t,j} = V_{a_{t,j}} + v_{t,j}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.19h}$$

$$\mathcal{E} = \sum_{j \in \mathcal{S}_T} p_{T,j} V_{T,j} \tag{4.2.19i}$$

$$\mathcal{R} = \sum_{j \in \mathcal{S}_T} p_{T,j} |V_{T,j} - \mathcal{E}| \tag{4.2.19j}$$

$$0 \leq s_{a_{t,j}} + X_{t,j} - s_{t,j} - w_{t,j} \leq d_{t,j}, \quad t \in \mathcal{T} j \in \mathcal{S}_t \tag{4.2.19k}$$

$$0 \leq s_{t,j} \leq \bar{s}, \quad t \in \mathcal{D}, j \in \mathcal{S}_t \tag{4.2.19l}$$

$$X_{T,j} \leq \bar{x}, \quad j \in \mathcal{S}_T \tag{4.2.19m}$$

$$u_{t,j,k} \in \mathbb{Z}_+, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, k \in \mathcal{F} \tag{4.2.19n}$$

$$s_{t,j} \in \mathbb{Z}_+, \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.19o}$$

The Pareto solutions of this problem are found by using an ϵ-constrained method. It is important to highlight that the SMMO problem does not seek to optimize the conditional expectation and risk at every time (as in traditional multi-stage SP formulations). Instead, the SMMO problem optimizes the *joint* expectation and risk (over the entire planning time). This formulation thus avoids ambiguity issues associated with time consistency of conditional risk evaluation encountered in traditional formulations. Another way to think about this difference is that our formulation first determines the accumulated cash flow over all stages and then optimizes its risk, while a traditional formulation determines the risk of the cash flow at each stage and then optimizes the accumulated risk over all stages.

4.2.4 *Multi-Product, Stochastic Setting*

We can conveniently extend the previous formulation to a multi-product setting. Assume that the investor now has a choice of producing multiple products $\mathcal{I} = \{i_1, i_2, \ldots, i_I\}$. We use $\alpha_{i,i'}, i, i' \in \mathcal{I}$ to represent the inter-dependencies between product i and i'. Specifically, $\alpha_{i,i'}$ denotes the units of product i' required to produce i. Note that $\alpha_{i,i} = 0, i \in \mathcal{I}$. For each product $i \in \mathcal{I}$, we define a set of technologies that can produce it; these technologies have capacities $\mathcal{B}^i$ and installation costs $\mathcal{C}^i$. Also, for each product $i \in \mathcal{I}$, we define a storage cost, waste disposal cost, operational cost, and selling price as ρ_s^i, ρ_w^i, ρ_p^i, and π_p^i. We also define a capacity limit for product i as $\bar{x}^i, i \in \mathcal{I}$, and storage limit for product i as $\bar{s}^i, i \in \mathcal{I}$. We define the demand for product i at stage t and scenario j as $d_{t,j}^i, t \in \mathcal{T}, j \in \mathcal{S}_t, i \in \mathcal{I}$. Our decision variables are the number of technologies with capacity B_k^i to be installed from the capacities list at stage t and scenario j for product i and these are modeled using the integer variables $u_{t,j,k}^i \in \mathbb{Z}_+, t \in \mathcal{D}, j \in \mathcal{S}_t, k \in \mathcal{F}^i, i \in \mathcal{I}$.

The total capacity installed at stage t and scenario j for product i is denoted as $x_{t,j}^i, t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I}$. The total installation cost at stage t and scenario j for product i is denoted as $y_{t,j}^i, t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I}$. These quantities are computed as:

$$x_{t,j}^i = \sum_{k \in \mathcal{F}^i} u_{t,j,k}^i B_k^i, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.20a}$$

$$y_{t,j}^i = \sum_{k \in \mathcal{F}^i} u_{t,j,k}^i C_k^i, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.20b}$$

The amount of storage at stage t and scenario j for product i is defined as $s_{t,j}^i \in \mathbb{Z}_+, t \in \mathcal{T}, j \in \mathcal{S}_t, i \in \mathcal{I}$, and the amount of product disposed at stage t and scenario j for product i is $w_{t,j}^i \in \mathbb{Z}_+, t \in \mathcal{T}, j \in \mathcal{S}_t, i \in \mathcal{I}$. The total production at stage t and scenario j for product i is denoted as $X_{t,j}^i, t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I}$. We also incorporate the cumulative installation cost occurred along the path to time t and scenario j for product i and denote

this as $Y^i_{t,j}, t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I}$. These quantities are computed as:

$$X^i_{t,j} = X^i_{a_{t,j}} + x^i_{a_{t,j}}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.21a}$$

$$Y^i_{t,j} = Y^i_{a_{t,j}} + y^i_{t,j}, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.21b}$$

The cost incurred at stage t and scenario j, is $q_{t,j}, j \in \mathcal{S}_T$ and is computed as:

$$q_{t,j} = \sum_{i \in \mathcal{I}} y^i_{t,j} + \rho^i_p X^i_{t,j} + \rho^i_s s^i_{t,j} + \rho^i_w w^i_{t,j}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t. \tag{4.2.22}$$

The profit incurred at stage t and scenario j is $r_{t,j}, t \in \mathcal{D}, j \in \mathcal{S}_t$ and is computed as:

$$r_{t,j} = \sum_{i \in \mathcal{I}} \pi^i_p (X^i_{t,j} + s^i_{a_{t,j}} - s^i_{t,j} - w^i_{t,j} - \sum_{i' \in \mathcal{I}} X^{i'}_{t,j} \alpha_{i',i}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t. \tag{4.2.23}$$

Under these definitions, we can define the rest of the quantities for cash flow, cumulative cash flow, and NPV in the same way that we did for the single-product case. This gives the SMMO problem:

$$\max_{u,s} \ \{\mathcal{E}, -\mathcal{R}\} \tag{4.2.24a}$$

$$\text{s.t. } x^i_{t,j} = \sum_{k \in \mathcal{F}^i} u^i_{t,j,k} B^i_k, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.24b}$$

$$y^i_{t,j} = \sum_{k \in \mathcal{F}^i} u^i_{t,j,k} C^i_k, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.24c}$$

$$X^i_{t,j} = X^i_{a_{t,j}} + x^i_{a_{t,j}}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.24d}$$

$$Y^i_{t,j} = Y^i_{a_{t,j}} + y^i_{t,j}, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.24e}$$

$$q_{t,j} = \sum_{i \in \mathcal{I}} y^i_{t,j} + \rho^i_p X^i_{t,j} + \rho^i_s s^i_{t,j} + \rho^i_w w^i_{t,j}, \quad t \in \mathcal{D}, j \in \mathcal{S}_t \tag{4.2.24f}$$

$$r_{t,j} = \sum_{i \in \mathcal{I}} \pi^i_p (X^i_{t,j} + s^i_{a_{t,j}} - s^i_{t,j} - w^i_{t,j} - \sum_{i' \in \mathcal{I}} X^{i'}_{t,j} \alpha_{i',i}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.24g}$$

$$v_{t,j} = \beta_t (r_{t,j} - q_{t,j}), \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.24h}$$

$$V_{t,j} = V_{a_{t,j}} + v_{t,j}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t \tag{4.2.24i}$$

$$\mathcal{E} = \sum_{j \in \mathcal{S}_T} p_{T,j} V_{T,j} \tag{4.2.24j}$$

$$\mathcal{R} = \sum_{j \in \mathcal{S}_T} p_{T,j} |V_{T,j} - \mathcal{E}| \tag{4.2.24k}$$

$$0 \leq s^i_{a_{t,j}} + X^i_{t,j} - s^i_{t,j} - w^i_{t,j} - \sum_{i \in \mathcal{I}} X^{i'}_{t,j} \alpha_{i',i} \leq d^i_{t,j}, \quad t \in \mathcal{T}, j \in \mathcal{S}_t, i, i' \in \mathcal{I} \tag{4.2.24l}$$

$$\alpha_{i,i'} X^i_{t,j} \leq s^{i'}_{a_{t,j}} + X^{i'}_{t,j}, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i, i' \in \mathcal{I} \tag{4.2.24m}$$

$$0 \leq s^i_{t,j} \leq \bar{s}^i, \quad t \in \mathcal{D}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.24n}$$

$$X^i_{T,j} \leq \bar{x}^i, \quad j \in \mathcal{S}_T, i \in \mathcal{I} \tag{4.2.24o}$$

$$\sum_{i \in \mathcal{I}} Y^i_{T,j} \leq \bar{y}, \quad j \in \mathcal{S}_T \tag{4.2.24p}$$

$$u^i_{t,j,k} \in \mathbb{Z}_+, \quad i \in \mathcal{D}, j \in \mathcal{S}_t, k \in \mathcal{F}^i, i \in \mathcal{I} \tag{4.2.24q}$$

$$s^i_{t,j} \in \mathbb{Z}_+, \quad t \in \mathcal{T}, j \in \mathcal{S}_t, i \in \mathcal{I} \tag{4.2.24r}$$

We highlight that the proposed formulation can be extended in a number of ways to add different investment logic (e.g., account for limited investment budgets). Here, we present a formulation that contains enough features to highlight benefits of modular technologies in mitigating risk.

4.3 Case Studies

In this section, we present different case studies to illustrate how modularization can help mitigate risk. The first case study involves a single-product setting with 3 stages and has a structure of a binary tree. We then present a more complex and realistic case study that includes interdependent products and more stages and scenarios. The optimization problems were solved using *Gurobi* (version 0.7.6) with a default MIP Gap of 0.01% and were implemented in the *JuMP* modeling framework. The scripts to reproduce all results can be found in https://github.com/zavalab/JuliaBox/tree/master/ModularPlanning.

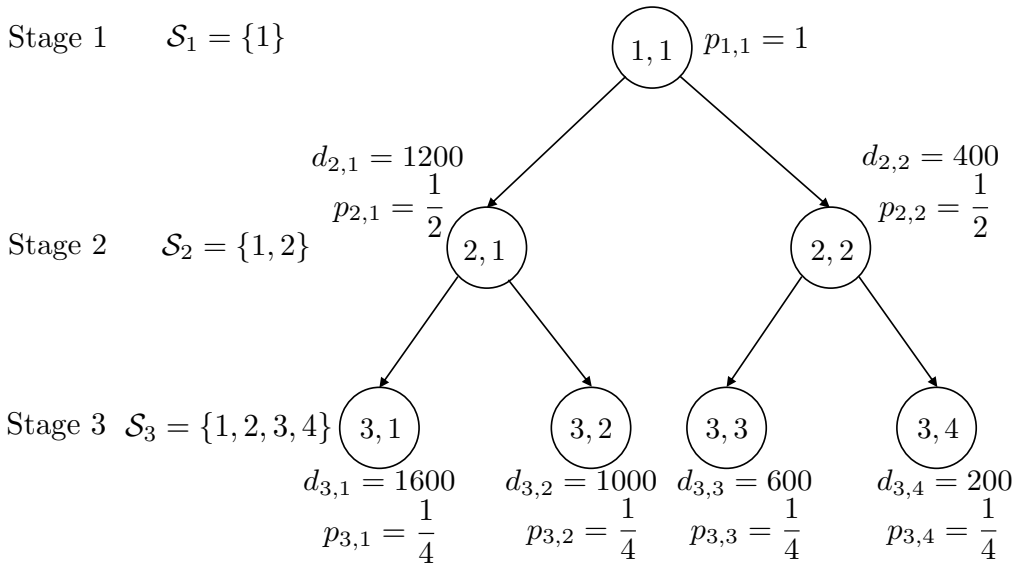


Figure 4.6: Tree representation for single-product stochastic case over 3 stages.

4.3.1 *Single-Product Problem*

Table 4.1: Data for single-product problem

Parameters	Values
Capacities (tons), $\mathcal{B}$	{100, 500, 1000, 1500}
Installation Cost ($), $\mathcal{C}$	{247, 721, 1145, 1500}
Capacity Limit (tons), $\bar{x}$	1500
Installation Cost Limit ($), $\bar{y}$	2000
Storage Cost ($), ρ_s	30
Storage Limit (tons), $\bar{s}$	400
Waste Cost ($), ρ_w	30
Operational Cost ($), ρ_p	50
Selling Price ($), π_p	140
Discount rate, γ	0.06

Figure 4.6 shows the stages, scenarios, and their corresponding demand and probabilities. The number inside each node represents the scenarios at each stage and the other

number next to the node indicates the demand for each scenario in tons. Each parent node has two children nodes and we assume that each outcome has equal probability. All other required data is summarized in Table 4.1. All the capacity-related quantities have the units of metric tons and price-related quantities have units of US dollars.

In this problem, we are seeking to make investment decisions at stage 1 and 2 that can help minimize NPV risk while achieving a constant level of expected NPV. To see the effect of that modular units have on flexibility, we solved this problem under three different capacity options (we call them Cases 1,2,3). In Case 1, we only allow the investor to choose between large capacities of 1500 tons and 1000 tons. In Case 2, we add medium capacity unit (500 tons) to the list to provide more flexibility. In Case 3, we allow the investor to choose from the complete capacity list (which includes smaller modular units). The three cases are solved for the undiscounted and discounted NPV problem (to see the impact of time value of money).

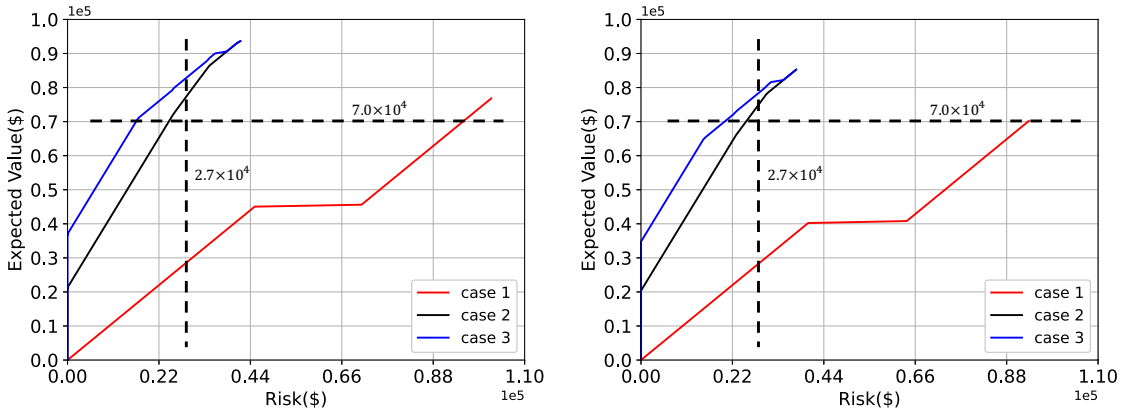


Figure 4.7: Pareto frontiers under undiscounted NPV (left) and discounted NPV (right) settings.

The Pareto frontiers for both problems are shown in Figure 4.7. Examples of investment plans obtained with these formulations are shown in Table 4.2. We can see that the shape of the Pareto frontiers for the discounted and undiscounted problems is similar. We can thus see that the discounting factor does not influence the decisions made at

each stage since the number of stages is small. As we will see in the next case study, the effect of discounting can be quite pronounced for problems that involve long planning horizons. The Pareto frontiers highlight that that a strong trade-off exists between the expected value and risk of the NPV (higher expected NPV results in higher risk). This trade-off arises from economies of scale and flexibility (it is less expensive but more risky to install large units). It is clear that the Pareto frontier for cases 2 and 3 (under which small units are available) dominate the frontier of case 1 (under which only large units are available). Importantly, this occurs even if the installation costs of the large units have better economies of scale. At the same level for the expected NPV, Cases 2 and 3 achieve a significantly lower risk (reduction by a factor of 3). Similarly, at the same risk level, Cases 2 and 3 achieve a much higher expected NPV (increase by a factor of 2). We can also see that Case 2 and 3 achieve levels of expected NPV that are not achievable in Case 1. Note that since we used ϵ-constrained method to obtain the Pareto solutions, the smoothness of the curve largely depends on the resolution of ϵ that was chosen. As different capacity plans result in discrete risk and expected value combinations, gaps (flat region of the line) may appear on the Pareto frontier.

In Table 4.2, the installation column shows installation decisions made for the undiscounted problem. These decisions are also visualized in Figures 4.8, 4.9, and 4.10. We can see that, with only larger capacity options (Case 1), we have no choice but to install the large unit at stage 1. In Case 2, we install a medium-sized unit in stage 1 and another medium-sized unit in stage 2; this achieves the same expected NPV but the risk is drastically reduced. For Case 3, we installed four small-sized units in the first stage and one medium-sized unit in the second stage. This achieves the same expected NPV but further decreases the risk. We thus conclude that the different capacity choices enable higher investment flexibility and reduced risk.

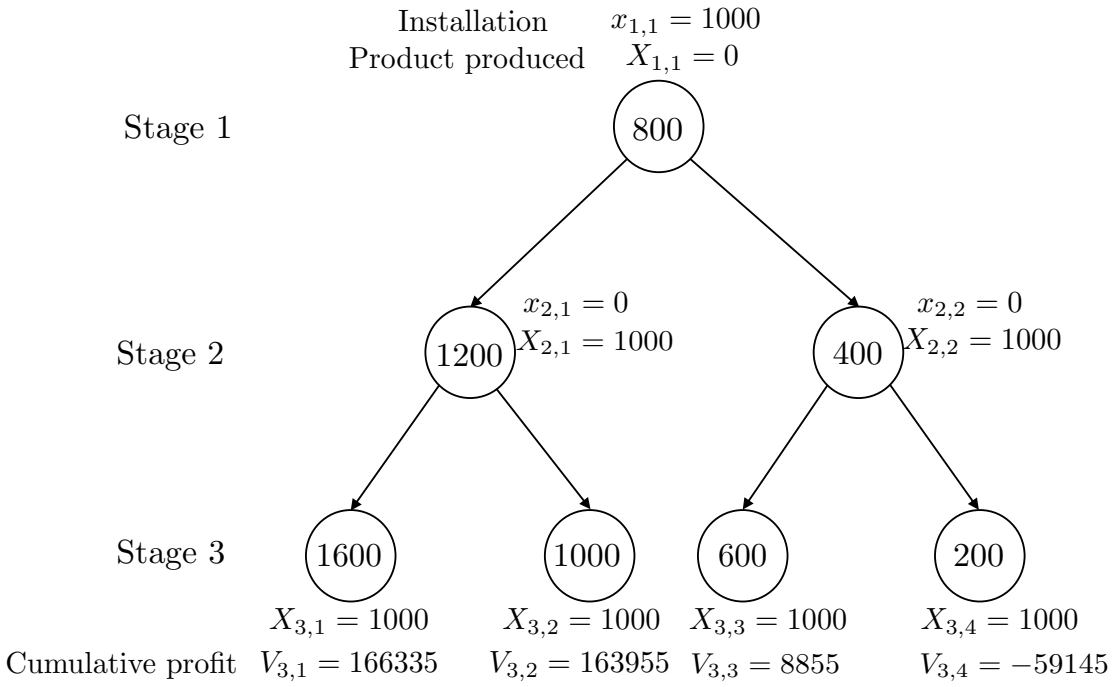


Figure 4.8: Investment strategy under Case 1 (undiscounted NPV setting).

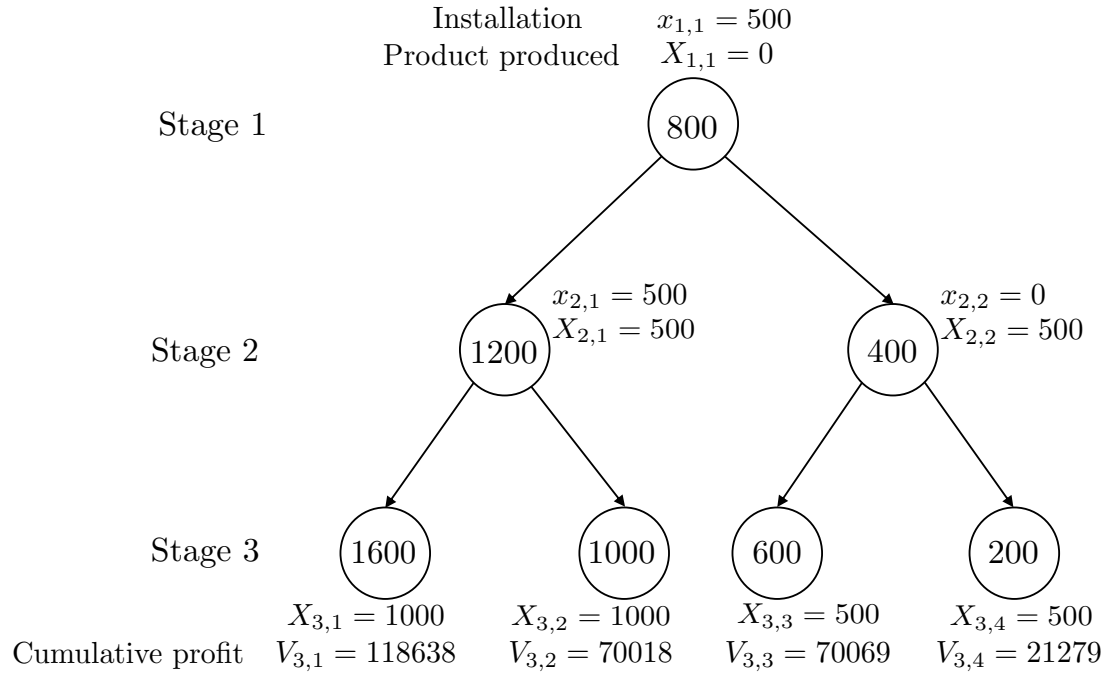


Figure 4.9: Investment strategy under Case 2 (undiscounted NPV setting).

Table 4.2: Investment strategy and associated risks (undiscounted NPV setting).

Cases	Technology Sizes (tons)	Risk ($)	Expected Value ($)
Case 1	{1000, 0, 0}	93149	7.0×10^4
Case 2	{500, 500, 0}	24361	7.0×10^4
Case 3	{100×4, 500 , 0}	16495	7.0×10^4

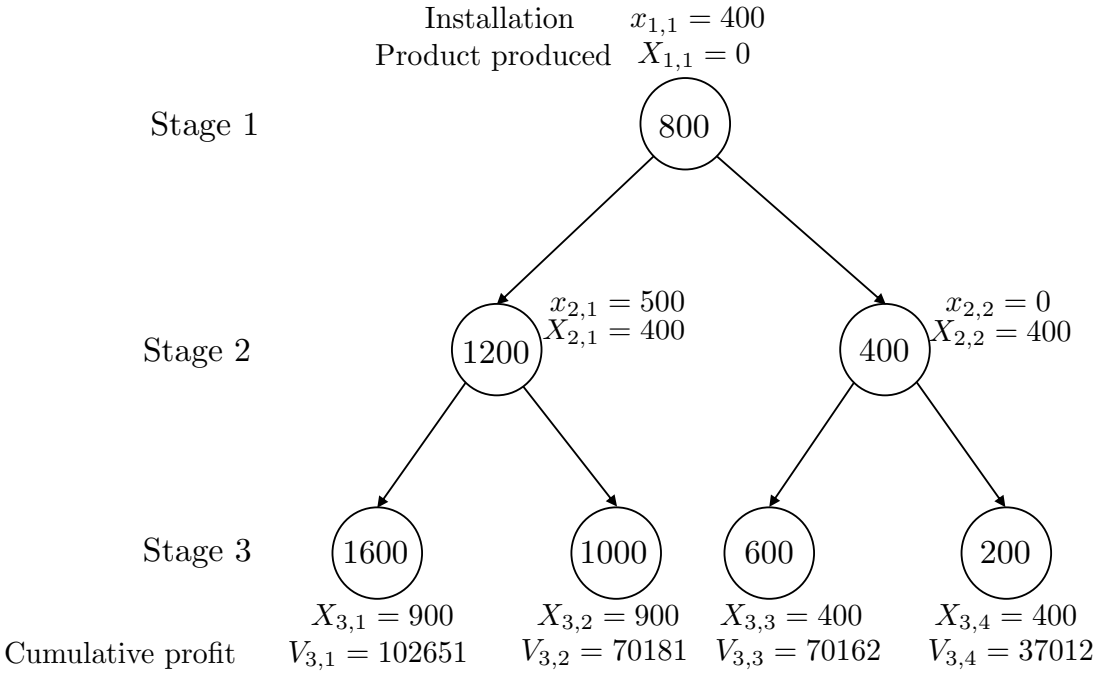


Figure 4.10: Investment strategy under Case 3 (undiscounted NPV setting).

4.3.2 *Multi-Product Problem*

Biogas is a methane-rich gas mixture that can be produced from anaerobic digestion of organic waste (such as cow manure). The biogas (in metric tons) can be sold directly or can be used as raw material to produce electricity (in MWh) and liquefied biomethane (in gallons). These products are represented as i_1, i_2, and i_3 respectively (Krich et al., 2005; Sampat et al., 2018). The process under study is visualized in Figure 4.11. The available technology capacities, investment cost, operation cost and other required information are summarized in Table 4.3 (Hu et al., 2018; Beddoes et al., 2007; Patel, 2019). The capital

Table 4.3: Data for biogas capacity expansion problem.

Parameters	Notation	Values
Capacities List	B^{i_1} (tons)	[400, 800, 1200]
	B^{i_2} (MWh)	[500, 1000, 2000]
	B^{i_3} (gallons)	[60000, 180000, 300000]
Installation Cost ($)	C^{i_1}	[272844, 374693, 457138]
	C^{i_2}	[172219, 234688, 347021]
	C^{i_3}	[577269, 1930312, 2935611]
Capacity Limit	$\bar{x}^{i_1}$ (tons)	6000
	$\bar{x}^{i_2}$ (MWh)	6000
	$\bar{x}^{i_3}$ (gallons)	1500000
Storage Cost	$\rho_s^{i_1}$ ($ per tons)	0
	$\rho_s^{i_2}$ ($ per MWh)	150000
	$\rho_s^{i_3}$ ($ per gallon)	0.2
Storage Limit	$\bar{s}^{i_1}$ (tons)	800
	$\bar{s}^{i_2}$ (MWh)	1000
	$\bar{s}^{i_3}$ (gallons)	180000
Waste Disposal Cost	$\rho_w^{i_1}$ ($ per ton)	0
	$\rho_w^{i_2}$ ($ per MWh)	0
	$\rho_w^{i_3}$ ($ per gallon)	2
Operational Cost	$\rho_p^{i_1}$ ($ per ton)	34
	$\rho_p^{i_2}$ ($ per MWh)	40
	$\rho_p^{i_3}$ ($ per gallon)	0.56
Selling Price	$\pi_p^{i_1}$ ($ per ton)	100
	$\pi_p^{i_2}$ ($ per MWh)	130
	$\pi_p^{i_3}$ ($ per gallon)	2.5
Interdependency	α_{i_2,i_1}	0.68
	α_{i_3,i_1}	0.0046
Total Installation Cost Limit ($)	$\bar{y}$	1×10^7
Interest Rate	γ	0.06

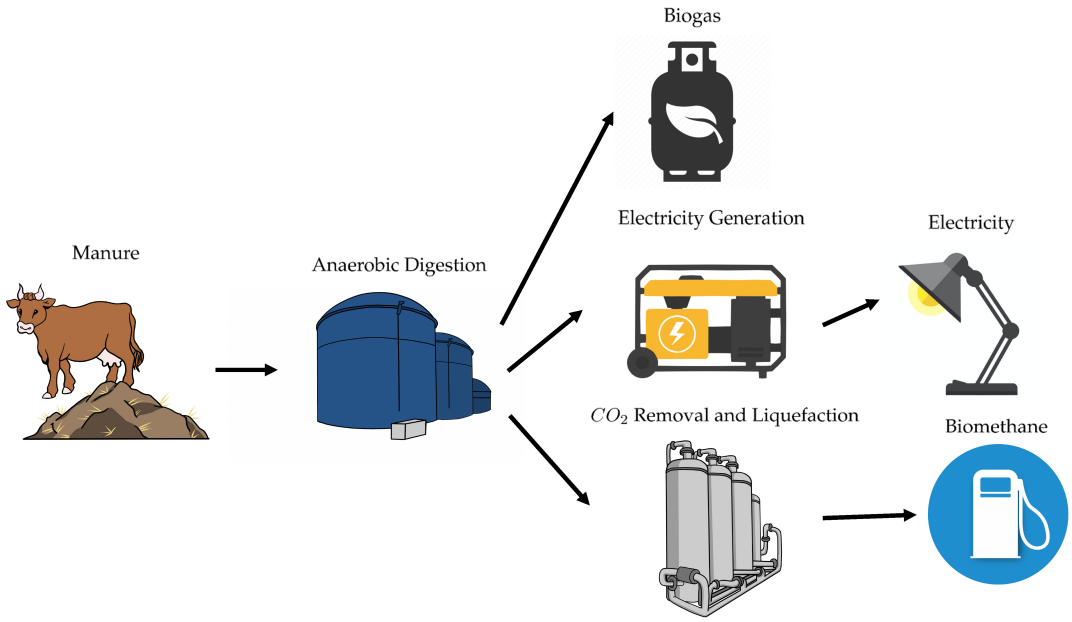


Figure 4.11: Process for the production of biogas and its byproducts

cost for these technologies roughly follows the 2/3 rule. The products are interdependent: producing 1 MWh of i_2 requires 0.68 tons of i_1 and producing 1 gallon of i_3 requires 0.0046 tons of i_1. The planning stages have a duration of one year; as such, all capacities and production levels are expressed on a per-year basis.

The stochastic multistage setting is illustrated in Figure 4.12. Here, we have a planning horizon with 10 stages. From stage 1 to stage 6, each parent node has two children nodes (which capture variability in market demands); after stage 6, the market is assumed to stay constant and thus each parent node only has one children node. The demand for selected nodes is shown next to the node. For the first six stages, the children nodes of each parent node represent an optimistic market and a pessimistic market.

We consider 3 possible cases; for Case 1, the unit for producing i_1 has a capacity of 1200 tons, the unit producing i_2 has a capacity of 2000 MWh, and the unit for producing i_3 has the capacity of 300000 gallons. In Case 2, we add a unit with a capacity of 800 tons

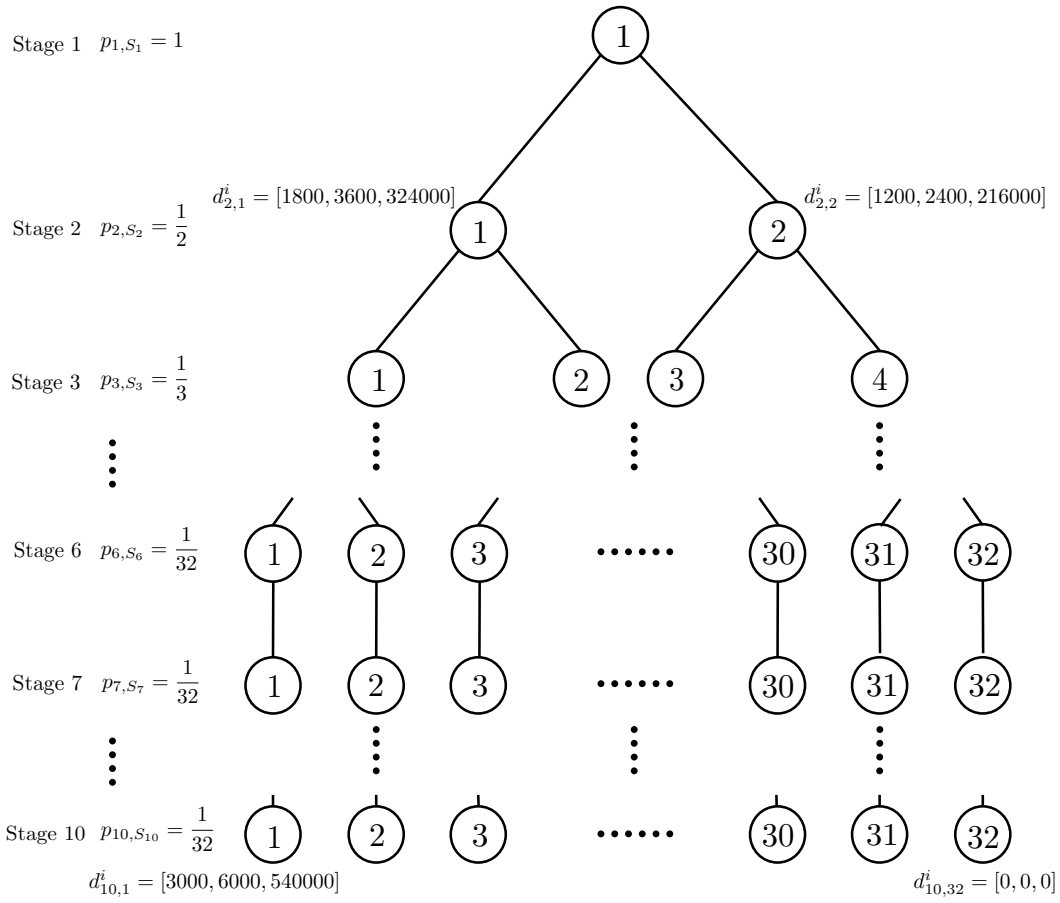


Figure 4.12: Tree representation of planning stages and scenarios of biogas case study.

for producing i_1, we add a unit with a capacity of 1000 MWh for the choices for i_2, and a capacity of 180,000 gallons for i_3. For Case 3, we further expand the capacity choices for product i_1 to include 400 tons, expand choices for i_2 to include 500 MWh, and expand choices for i_3 to include 60,000 gallons. To provide some context on the size of these units, an annual capacity of 500 MWh corresponds to a power capacity of 500/8760=0.057 MW (57 kW). As such, the small capacities for the power generators correspond to those of small modular systems.

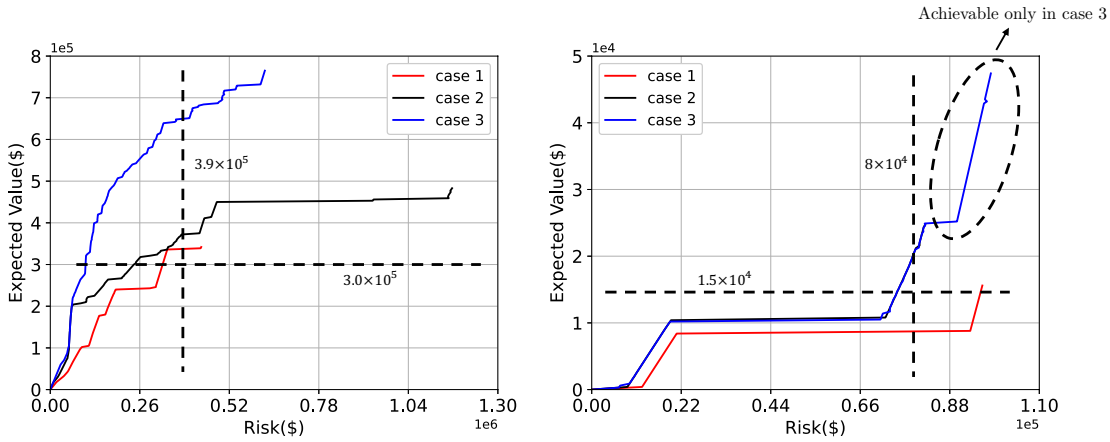


Figure 4.13: Pareto frontiers for undiscounted NPV (left) and discounted NPV (right) settings.

We would like to determine the number of planning stages that it takes for the investment to be profitable. As such, we gradually increase the planning horizon of the CE problem until the profit is positive. We found that, for both discounted and undiscounted problems, the expected NPV remains zero for any planning horizon with less than 8 stages. In other words, the investment is only profitable if the project lifetime is at least 8 years. We can thus see that the length of the planning horizon plays an important role in making investment decisions. We assume that the planning horizon is 10 years (as shown in Figure 4.12). Again, we would like to determine an optimal investment strategy that maximizes expected NPV and minimizes its risk. The Pareto frontiers are shown in Figure 4.13 and we compare risks obtained under the different cases in Table 4.4 and Table 4.6.

Table 4.4: Investment strategy for undiscounted NPV problem with $\mathcal{E} = 3.0 \times 10^5$

Cases	Installation	Expected Value ($)	Risk ($)	# of Constraints	# of Variables (Cont.+Int.)
Case 1	$x^{i_1}_{1,1} = 1200 \times 2$ tons $x^{i_2}_{1,1} = 2000$ MWh	3.0×10^5	3.26×10^5	4355	2225+1623
Case 2	$x^{i_1}_{1,1} = 1200$ tons $x^{i_2}_{1,1} = 1000$ MWh $x^{i_1}_{2,1} = 1200$ tons $x^{i_3}_{2,1} = 180000$ gallons $x^{i_1}_{3,1} = 1200$ tons $x^{i_2}_{5,1} = 2000$ MWh $x^{i_2}_{5,4} = 1000$ MWh $x^{i_2}_{6,6} = 2000$ MWh	3.0×10^5	2.43×10^5	4355	2225+2100
Case 3	$x^{i_1}_{1,1} = 800$ tons $x^{i_3}_{1,1} = 60000$ gallons $x^{i_1}_{3,1} = 1200 \times 2$ tons $x^{i_3}_{3,1} = 60000 \times 2$ gallons	3.0×10^5	1.02×10^5	4355	2225+2577

We again find that the Pareto frontier of Case 3 (considering small technologies) dominates. Looking horizontally (for the same expected NPV) cases with more capacity options reduce risk. Looking vertically (for the same risk) we can see that modularity allows us to reach higher expected NPVs. For the discounted NPV problem we see that adding smaller capacity options (Case 2) reduces risk, further reducing the capacity (Case 3) can achieve higher expected profits but does not help to mitigate the risk. This is because of complex interplays between discounting and economies of scale. As we discount the future cash flow, the effect of installing small capacities at future stages reduces, and together with the effect of economies of scale, the advantages brought by modular technologies become less obvious. This indicates that reducing technology sizes aids flexibility (but there is a limit to such flexibility). From Table 4.4 and Table 4.5 we can see that,

Table 4.5: Investment strategy for discounted NPV problem with $\mathcal{R} = 3.9 \times 10^5$

Cases	Installation	Expected Value($)	Risk ($)
Case 1	$x_{1,1}^{i_1} = 1200 \times 2$ tons $x_{1,1}^{i_2} = 2000$ MWh	2.64×10^5	3.90×10^5
Case 2	$x_{1,1}^{i_1} = 1200$ tons $x_{2,1}^{i_1} = 1200 \times 2$ tons $x_{2,1}^{i_2} = 2000$ MWh $x_{2,1}^{i_3} = 180000$ gallons $x_{3,4}^{i_2} = 1000$ MWh $x_{6,24}^{i_2} = 1000$ MWh	3.72×10^5	3.90×10^5
Case 3	$x_{1,1}^{i_1} = 1200$ tons $x_{1,1}^{i_3} = 60000 \times 2$ gallons $x_{2,1}^{i_1} = 1200 \times 2$ tons $x_{2,1}^{i_2} = 2000$ MWh $x_{3,1}^{i_3} = 60000$ tons	6.50×10^5	3.90×10^5

Table 4.6: Investment strategy for discounted NPV problem with $\mathcal{E} = 1.5 \times 10^4$

Cases	Installation	Expected Value($)	Risk ($)
Case 1	$x_{1,1}^{i_1} = 1200$ tons	1.5×10^4	3.06×10^6
Case 2	$x_{1,1}^{i_1} = 1200$ tons $x_{3,4}^{i_2} = 1000$ MWh $x_{5,12}^{i_2} = 1000$ MWh $x_{6,16}^{i_2} = 1000$ MWh	1.5×10^4	2.41×10^6
Case 3	$x_{1,1}^{i_1} = 1200$ tons $x_{3,4}^{i_2} = 1000$ MWh $x_{5,12}^{i_2} = 1000$ MWh $x_{6,16}^{i_2} = 1000$ MWh	1.5×10^4	2.41×10^6

for the undiscounted problem, most of the investment occurs at the early stages. Here, we can also see that modular technologies are used extensively to reduce risk (risk is reduced by a factor of three) and increase profit (profit is increased by a factor of three).

4.4 Conclusion and Future Work

We study logistical investment flexibility provided by modular processing technologies for mitigating risk. Specifically, we propose a capacity expansion problem that aims to determine optimal investment strategies over a given planning horizon. This expansion problem is a stochastic, multi-stage, and multi-objective optimization problem. The formulation accounts for multi-product dependencies between small/large units and for trade-offs between expected profit and risk. The formulation uses a cumulative risk measure to avoid time-consistency issues of traditional, per-stage risk-minimization formulations and we argue that this approach is more compatible with typical investment metrics such as the net present value. Case studies of different complexity are presented to illustrate the developments. Our studies reveal that the Pareto frontier of a flexible setting (allowing for deployment of small units) dominates the Pareto frontier of an inflexible setting (allowing only for deployment of large units). Notably, this dominance is prevalent despite benefits arising from economies of scale of large processing units. Small technologies provide flexibility that translates into tangible reductions of risk (despite the fact that they are not benefited by economies of scale). However, we also find that flexibility provided by capacity reductions has limits that result from the complex interplay between economies of scale and discounting.

As part of future work, we are interested in exploring the use of decomposition strategies to address tractability issues (e.g., by using stochastic dual dynamic programming techniques). In this work we ignored engineering costs associated with different types of technologies (which can be reduced using modularization). We will use more detailed cost representations and case studies in future work.

5

A SPATIAL SUPERSTRUCTURE APPROACH TO THE OPTIMAL DESIGN OF MODULAR SYSTEMS

This article is submitted to Computers and Chemical Engineering and is currently under review.

5.1 Introduction

Modularity is a design principle that aims to provide flexibility for spatio-temporal assembly, disassembly, and reconfiguration of systems. This design principle can be applied to multi-scale manufacturing systems that connect equipment units/technologies, processes (collections of units/technologies), facilities (collections of processes), and entire supply chains (collections of facilities). Modularity principles have been recently explored in diverse industrial sectors such as power generation, data centers, and chemical processes (Frivaldsky et al., 2018; Berthélemy and Rangel, 2015; Dong et al., 2009; Chakraborty et al., 2009; R., 1999). It is a design principle that can be applied at different levels of an organization; for instance, processes that compose a facility (collection of processes at a given geographical location) can be interpreted as modules and products exchanged between such modules give rise to a facility. The connectivity induced from products exchanged

between processes and from product transformation in such processes induces a degree of modularity of the facility. For instance, facilities that have dense product interdependencies are less modular than those that have sparse interdependencies. The degree of product interdependency affects flexibility, as facilities that are tightly coupled are typically more difficult to reconfigure. Similarly, at a higher organization level, facilities that compose a supply chain can be interpreted as modules that exchange products across geographical locations (e.g., via long-distance transport). This indicates that a supply chain can be seen as a distributed network of processes (a distributed facility with processes placed at different geographical locations), while a typical facility can be seen as a centralized network of processes (all processes are placed at the same geographical location). As in the case of a facility, the modularity of a supply chain is affected by the connectivity induced from product transport across components, from product transformation in its components, and from its ability to be reconfigured (e.g., movable processes). For instance, a supply chain composed of small processes (easier to move reconfigure) is more flexible than one composed of large processes (difficult to move and reconfigure) (Zhao et al., 2018). Similarly, a supply chain with sparse product connectivity will be more modular than that with dense product connectivity. This is because, when a process fails in a densely connected supply chain, it can trigger a collapse of the entire system.

Modular design studies reported in the literature have focused mostly on single processes and thus do not assess how modularization can help mitigate system-wide (i.e., at facilities and supply chain levels). Recent work in power grid and natural gas networks has revealed that deploying distributed data centers, batteries, vehicle charging stations, gas-fired power plants, and manufacturing facilities can add flexibility, relieve network congestion, and enhance system-wide performance (Kim et al., 2017b; Sioshansi et al., 2009; Chiang and Zavala, 2016). This flexibility can be used to absorb fluctuations of wind and solar power and can help withstand externalities (e.g., extreme weather events, policy, equipment failures). One could thus argue that widespread deployment of small modular systems can, in principle, provide network flexibility to mitigate system-

wide risks (e.g., climate change). Modular supply chain design that considers demand uncertainty to manage risk has recently been explored. A mixed integer stochastic programming problem is formulated to determine the number of modules, the locations of facilities and the flow of materials to minimize the downside risk and the overall profit at the same time (Bhosekar et al., 2021). However, this work considers modularity only at the facility level, instead of at the system level that involves the entire supply chain.

Maximal p-graph structures and superstructures are system representations that have been widely used for the design of chemical processes and of mass/energy recovery networks for single facilities (F Friedler and Fan, 1992; Christodoulos A Floudas and Grossmann, 1986). A maximal p-graph structure encodes all possible feasible paths between primary products, processing tasks, and intermediate and final products (Yeomans and Grossmann, 1999). A superstructure encodes all possible configurations of equipment units and product flows that perform tasks defined by the maximal p-graph (i.e., multiple units might perform the same task) (El-Halwagi and Manousiouthakis, 1989; Isafiade and Fraser, 2009). While these representations provide a powerful framework to investigate systems at a process level, they do not encode spatial information, which is necessary to capture how design affects flexibility at higher organization levels (facilities and supply chains).

In this work, we propose an optimization framework to facilitate the design of modular processes, facilities, and supply chains. Central to our approach is the concept a spatial superstructure, which is a graph that encodes all possible dependencies between components. We show that the spatial superstructure is a generalization of the superstructure and p-graph used for process design in that it encodes spatial (geographical) context. Moreover, we show that this generalization enables the simultaneous design of processes, facilities, and of supply chains in a unified manner. Specifically, a spatial superstructure is a superstructure under which technologies and flows encode positional (geographical) context and that encodes product dependencies that arise from transformation (as in a p-graph). This allows us to represent standard centralized processes and

facilities (under which technologies are placed at the same geographical location) and a spatially-distributed process (under which technologies are distributed over multiple geographical locations) by using the same graph topology. The graph representation reveals that a key distinction between a spatial superstructure and other representations is in how product transportation is accounted for. For instance, short-range transport (inside a process or facility) might use pipelines while long-range transport might use truck hauling or railways.

The proposed approach leverages the graph representation of the spatial superstructure to identify topologies that minimize system design cost and that maximize design modularity. We show that this design problem can be cast as a mixed-integer, multi-objective optimization formulation and allows us to capture interdependencies between primary products (raw materials), intermediate products, and final products that arise from product transformation and transport across components. We also leverage the topology of the spatial superstructure to accelerate the optimal design search by restricting such search along feasible paths that obtain desired products from primary products. This approach contrasts with standard superstructure optimization approaches that search over individual technologies/units. We demonstrate the capabilities for the design of a plastic waste upcycling supply chain.

5.2 Concepts and Graph Representations

In this section, we revisit the concept of a p-graph, maximal p-graph, and superstructure and provide a unifying graph-theoretic perspective. We use these concepts to propose a spatial superstructure that will be used to guide the design of modular systems.

5.2.1 *P-graph and Maximal p-graph*

In the context of chemical processes, graphs have been used to analyze interdependencies between products and technologies (unit operations) in a process and with this unravel a

number of fundamental systems properties such as topological feasibility (e.g., ability to reach a set of products from a set of primary products).

A process can be represented/modeled as a p-graph (short for *process graph*). In a p-graph, nodes represent technologies (processing tasks or unit operations) and products (primary products as well as intermediate and final products) while edges represent dependencies between products and technologies. Here, the concept of product is general in that it can capture general resources such as energy (e.g., electricity). In addition, we note that technologies induce complex interdependencies because they conduct transformation of products into other products (e.g., a chemical reactor or a separation unit). Under the p-graph abstraction, it is possible to derive a *maximal p-graph* (max p-graph for short) that encodes all possible technologies and required primary products and intermediate products that can be used to obtain reach a desired set of final products. This representation is powerful and insightful because any possible process configuration that connects primary products, technologies, and intermediate/final products is embedded in the maximal p-graph. A specific process realization is derived by selection of specific nodes and edges (which form a path between primary products (i.e., primary products), intermediate products, and final products desired). As we will see, superstructure representations inherit the topology of max p-graphs.

Suppose that a process involves a set of intermediate/final products $\mathcal{P}$ and a set of primary products $\mathcal{R}$ from which intermediate/final products are derived (via technologies). Furthermore, we define a set of all products involved in the process as $\mathcal{I}$. A product can potentially be generated by different types of technologies (techs for short), and we define a set of possible techs as $\mathcal{T}$.

Associated with each tech $t \in \mathcal{T}$, there is a set of output products $\Omega_t \in \mathcal{P}$, a set of input products $\mathcal{K}_t \in \mathcal{I}$, and a tech type θ_t. For convenience, we categorize techs by products and types using the subsets $\mathcal{T}_{i,i',j} \subseteq \mathcal{T}$ with $\mathcal{T}_{i,i',j} := \{t | i \in \Omega_t, i' \in \mathcal{K}_t, \theta_t = j\}$.

We model the *max p-graph* as a directed graph $\mathcal{G}^p = (\mathcal{N}^p; \mathcal{E}^p)$ where $\mathcal{N}^p$ is its set of nodes (vertices) and $\mathcal{E}^p$ is its set of edges. The set of nodes include product nodes and tech

nodes. We define the set of nodes representing the supplies/sources of primary product as $\mathcal{S}^p \subseteq \mathcal{N}^p$). Associated with each node $s \in \mathcal{S}^p$ there is a type of primary product $\Omega_s \in \mathcal{R}$. For convenience, we categorize suppliers as $\mathcal{S}_i^p \subseteq \mathcal{S}^p$ with $\mathcal{S}_i^p := \{s|\Omega_s = i\}$. Similarly, we define the set of nodes representing demands/sinks of final products as $\mathcal{D}^p \subseteq \mathcal{N}^p$. Associated with each $d \in \mathcal{D}^p$, there is a type of product $\mathcal{K}_d \in \mathcal{P}$; we categorize demand nodes as $\mathcal{D}_i^p \subseteq \mathcal{D}^p$ with $\mathcal{D}_i^p := \{d|\mathcal{K}_d = i\}$.

We define the set of nodes representing the techs as $\mathcal{U}^p \subseteq \mathcal{N}^p$. For each node $u \in \mathcal{U}^p$, there is a tech $\tau_u \in \mathcal{T}$ associated with it, and we classify tech nodes as $\mathcal{U}_t^p \subseteq \mathcal{U}^p$ with $\mathcal{U}_t^p := \{u|\tau_u = t\}$. Because each tech t is also associated with an product sets $\Omega_t, \mathcal{K}_t$ and a type θ_t, we have that each node $u \in \mathcal{U}^p$ is associated with a set of products $\Omega_{\tau_u} \in \mathcal{P}$ that the tech generates, a set of products $\mathcal{K}_{\tau_u} \in \mathcal{I}$ that enter the tech, and a type of tech θ_{τ_u}. For convenience, we use the short-hand notation Ω_u, $\mathcal{K}_u$, and θ_u. We define the subsets $\mathcal{U}_{i,i',j}^p := \{u|i \in \Omega_u, i' \in \mathcal{K}_u, \theta_u = j\}$. We highlight that, in the max p-graph representation, the tech node set $\mathcal{U}^p$ contains only attributes of tech $t \in \mathcal{T}$, so they are defined similarly. In other words, node $\mathcal{U}_t^p = \mathcal{U}_{i,i',j}^p$ corresponds to the tech $t = \mathcal{T}_{i,i',j}$. Finally, the set of all nodes is:

$$\mathcal{N}^p = \mathcal{S}^p \cup \mathcal{D}^p \cup \mathcal{U}^p. \tag{5.2.1}$$

Figure 5.1 provides an illustration of a max p-graph and showcases how complex interdependencies between products and techs arise. In this example, the set $\mathcal{P}$ contains products i_3 and i_5, set $\mathcal{R}$ contains primary products i_1, i_2, i_3, i_4 and i_5, and the set $\mathcal{I}$ contains the final product i_5. The intermediate product i_3 and final product i_5 are also included in the primary products because we consider the possibility of satisfying the demand by purchasing it from an external market. The set $\mathcal{T}$ contains a couple of tech types producing product i_3 from either i_1 or i_1 and i_2, and 2 types of techs producing product i_5 from either i_3 or i_3 and i_4, represented as $\{\mathcal{T}_{i_3,i_1,j_1}, \mathcal{T}_{i_3,\{i_1,i_2\},j_2}, \mathcal{T}_{i_5,\{i_3,i_4\},j_1}, \mathcal{T}_{i_5,i_3,j_2}\}$. Nodes representing supplies of primary products are on the left, shown as nodes $S_{i_1}^p$, $S_{i_2}^p$,

$S^p_{i_3}$, $S^p_{i_4}$, and $S^p_{i_5}$ in set $\mathcal{S}^p$; nodes representing technologies are in the middle, shown as node $U^p_{i_3,i_1,j_1}$, $U^p_{i_3,\{i_1,i_2\},j_2}$, $U^p_{i_5,\{i_3,i_4\},j_1}$ and $U^p_{i_5,i_3,j_2}$ in set $\mathcal{U}^p$; nodes representing the demand are on the right, shown as node $D^p_{i_5}$ in set $\mathcal{D}^p$. Nodes and edges are highlighted based on the associated product; for example, node $S^p_{i_4}$ and the edge that carries the product i_4 to technology $U^p_{i_5,\{i_3,i_4\},j_1}$ have the same color. The product hierarchy of the process is also displayed in this max p-graph representation; looking from top to bottom, we can see that products or techs that are involved in the early stage of the process are on the top and those that are involved later in the process are on the bottom. Moreover, the max p-graph tells us that primary products i_1, i_2, and i_4 can only be purchased but not produced. Product i_3 is an intermediate product that can be produced by either type of tech or purchased from the external market; this product is also fed to the techs that produce product i_5 (which is the final product).

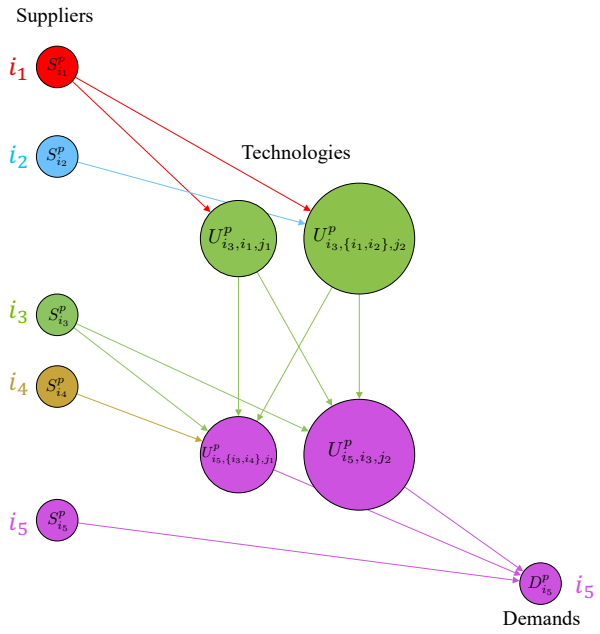


Figure 5.1: Illustration of a max p-graph showing dependencies between products and technologies.

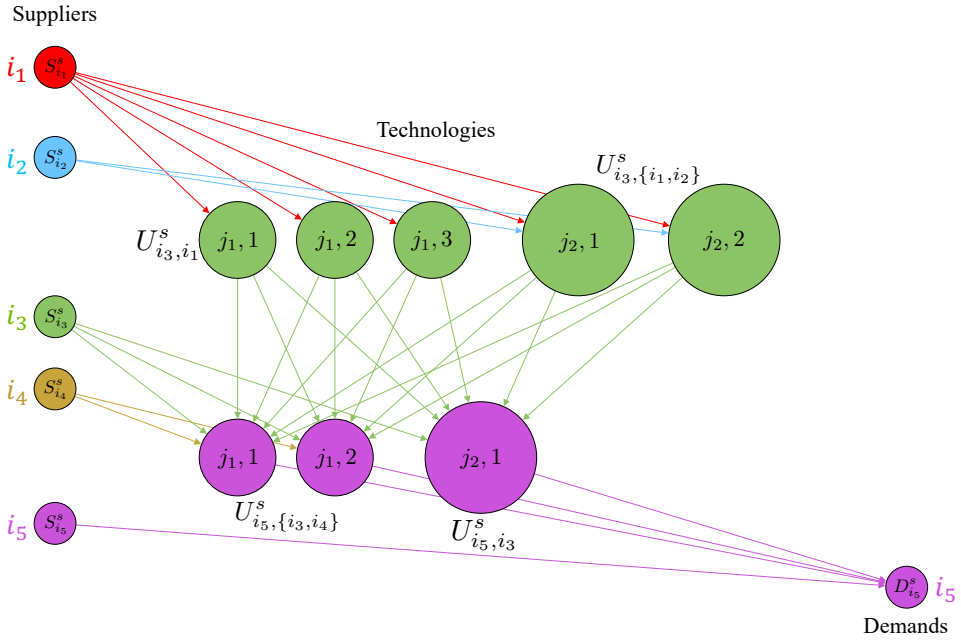


Figure 5.2: Illustration of a superstructure (associated with max p-graph in Figure 5.1) showing dependencies between products and technologies.

5.2.2 *Superstructure*

A superstructure is a system representation that inherits properties of a max p-graph (product-tech connectivity) but also accounts for the possibility of having multiple units/copies of techs and also account for additional attributes of techs (e.g., capacities). It is thus important to highlight that a superstructure can be derived from the topology of a max p-graph. The key difference is that, in a max p-graph, techs are interpreted as unit operations while, in a superstructure, techs represent equipment units.

A superstructure can also be represented as a graph; specifically, it can be represented as a directed graph $\mathcal{G}^s = (\mathcal{N}^s; \mathcal{E}^s)$, where $\mathcal{N}^s$ is its set of nodes (vertices) and $\mathcal{E}^s$ is its set of edges. As in the p-graph, primary products are defined as $\mathcal{R}$; intermediate/final products are defined as $\mathcal{P}$; and all products are defined as $\mathcal{I}$; techs are defined as $\mathcal{T}$ with the same attributes categories as in the max-p graph. Nodes representing supplies and demands are defined as $\mathcal{S}^s$ and $\mathcal{D}^s$, and the set of nodes $\mathcal{U}^s$ that represents techs. In this representation, each tech $u \in \mathcal{U}^s$ has an additional attribute that represents the unit number η_u.

We categorize the nodes $\mathcal{U}^s$ as $\mathcal{U}^s_{t,h} = \mathcal{U}^s_{i,i',j,h} \subseteq \mathcal{U}^s$ with $\mathcal{U}^s_{t,h} := \{u | \tau_u = t, \eta_u = h\}$ and $\mathcal{U}^s_{i,i',j,h} := \{u | i \in \Omega_u, i' \in \mathcal{K}_u, \theta_u = j, \eta_u = h\}$. In the case of a superstructure, the set $\mathcal{U}^s$ adds another layer of information on top of the set $\mathcal{T}$ to indicate that multiple copies of the same tech might be available. In other words, node $\mathcal{U}^s_{t,h} = \mathcal{U}^s_{i,i',j,h}$ corresponds to the h unit/copy of tech $\mathcal{T}_{i,i',j}$. The set $\mathcal{N}^s$ for all nodes in the superstructure is $\mathcal{N}^s = \mathcal{S}^s \cup \mathcal{D}^s \cup \mathcal{U}^s$.

Using the same example shown in Figure 5.1, we illustrate a superstructure representation of this system in Figure 5.2. Here, we labeled the set $\mathcal{U}^s_{i,i'} := \{u | i \in \Omega_u, i' \in \mathcal{K}_u\}$ on the side and labeled the attribute $\{\theta_u, \eta_u\}$ on the node for each $u \in \mathcal{U}^s$. Comparing to the max p-graph representation, only the notation for nodes representing techs has changed. For instance, the notation $U^s_{i_3,i_1,j_1,1}$ denotes the first unit/copy of tech $\mathcal{T}_{i_3,i_1,j_1}$ that is used to obtain the intermediate product i_3. Multiple units/copies of this tech are available to satisfy the demand of final product i_5 and the same applied for other techs. We can also

observe that the product-tech connectivity of the max p-graph is inherited by the superstructure. Moreover, we see that the superstructure graph is much denser than that of the max p-graph (due to the availability of multiple tech units).

5.2.3 *Spatial Superstructure*

We now proceed to generalize the notion of a max p-graph and of a superstructure to capture spatial context. Capturing spatial context is necessary to represent flexibility provided by modularity at different scales (process, facilities, supply chains). The *key observation* is that a supply chain can be seen as a distributed facility that exchanges products between processes (placed at different geographical locations). Similarly, a centralized facility can be seen as a supply chain with a single geographical location. This *unifying view* of a system will reveal interesting insights that can be exploited to derive a general graph-theoretic framework that explains how modularity emerges in a system design. Specifically, we will see that the topology of the spatial superstructure directly inherits the topology of a superstructure, which in turn inherits the product-tech topology of a max p-graph. Exploiting this topological dependencies is key in building superstructures and in identifying feasible system designs. Our final aim will be to derive optimization formulations identify a subgraph from the spatial superstructure (a design) to obtain a supply chain (composed of processes and facilities of different sizes and at potentially multiple locations) that minimizes system-wide cost and that maximizes modularity.

A spatial superstructure is a superstructure under which techs and connections encode positional context. This allows us to represent a standard single-site process/facility (under which equipment units are located at the same geographical location) and a spatially-distributed process/facility (under which units are distributed over multiple geographical locations) by using the same graph topology. The spatial superstructure allows us to capture transportation modes for the products and associated constraints and costs; for instance, short-range transport (inside a location) might use pipelines while

long-range transport (across locations) might use trucks or railways.

The graph representation of the spatial superstructure is inherited from that of the superstructure. We model the spatial superstructure as a directed graph $\mathcal{G}^q = (\mathcal{N}^q; \mathcal{E}^q)$ where $\mathcal{N}^q$ is its set of nodes (vertices) and $\mathcal{E}^q$ is its set of edges. We define a set of potential spatial locations for placing technologies as $\mathcal{G}_t$, a set of potential locations for suppliers as $\mathcal{G}_s$, and a set of potential locations for demands as $\mathcal{G}_d$. We then define a set of all locations as $\mathcal{G} = \mathcal{G}_t \cup \mathcal{G}_s \cup \mathcal{G}_d$.

As in the max p-graph representation, primary products are defined as $\mathcal{R}$; intermediate and final products are defined as $\mathcal{P}$; all products are defined as $\mathcal{I}$; techs are defined as $\mathcal{T}$ with the same attributes and nested representation. Nodes representing supplies are defined as $\mathcal{S}^q$, and for nodes $s \in \mathcal{S}$, there is a new attribute $\phi_s \in \mathcal{G}_s$ representing the location of the supplies/sources of primary products. We define subsets to categorize suppliers by location and product as $\mathcal{S}^q_{i,g} \subseteq \mathcal{S}^q_i \subseteq \mathcal{S}^q$ with $\mathcal{S}^q_{i,g} := \{s | \Omega_s = i, \phi_s = g\}$ and $\mathcal{S}^q_i := \{s | \Omega_s = i\}$. Nodes representing demands are defined as $\mathcal{D}^q$, and for node $d \in \mathcal{D}^q$, there is a new location attribute $\phi_d \in \mathcal{G}_d$. We define the categorization as subsets $\mathcal{D}^q_{i',g} \subseteq \mathcal{D}^q_{i'} \subseteq \mathcal{D}^q$ with $\mathcal{D}^q_{i',g} := \{d | \mathcal{K}_d = i', \phi_d = g\}$ and $\mathcal{D}^q_{i'} := \{d | \mathcal{K}_d = i'\}$.

For each tech node $u \in \mathcal{U}^q$ there is a new location attribute $\phi_u \in \mathcal{G}_t$, and the subsets are written as $\mathcal{U}^q_{t,h,g} = \mathcal{U}^q_{i,i',j,h,g} \subseteq \mathcal{U}^q$ with $\mathcal{U}^q_{t,h,g} := \{u | \tau_u = t, \eta_u = h, \phi_u = g\}$ and $\mathcal{U}^q_{i,i',j,h,g} := \{u | i \in \Omega_u, i' \in \mathcal{K}_u, \theta_u = j, \eta_u = h, \phi_u = g\}$. Specifically, node $\mathcal{U}^q_{t,h,g} = \mathcal{U}^q_{i,i',j,h,g}$ corresponds to the h copy of tech $\mathcal{T}_{i,i',j}$ that is located at location g. The set of all nodes is $\mathcal{N}^q = \mathcal{S}^q \cup \mathcal{D}^q \cup \mathcal{U}^q$.

Using the same example shown in Figure 5.2 and a couple of potential locations A and B, we illustrate the spatial superstructure in Figure 5.3. Here, we labeled the set $\mathcal{U}^q_{i,i'}$ for the nodes on the side while labeled the attribute $\{\theta_u, \eta_u, \phi_u\}$ on the nodes. We note that there might be multiple locations for suppliers and demands. We also note that edges that connect nodes at the same locations or across different locations have different meaning. For example, we can choose to install the first copy of tech $\mathcal{T}_{i_3,i_1,j_1}$ that produces i_3 at location A (represented as node $U^q_{i_3,i_1,j_1,1,A}$) and the first copy of tech $\mathcal{T}_{i_5,i_3,j_2}$ that

produces i_5 at location B (represented as node $U^q_{i_5,i_3,j_2,1,B}$), and the edge connecting them represents the transportation of product i_3 from location A to location B. If these techs are both placed at location A, an edge connecting them represents short-range (on-site) transport. We observe that the topology of the spatial superstructure is inherited from that of the superstructure, which in turn inherits the product-tech connectivity from the max p-graph. We also note that the spatial superstructure is much *larger and denser* than the superstructure and can becomes difficult (if not impossible) to express and visualize, due to the potential need to capture many geographical locations and tech units at such locations. Therefore, deriving an automatic approach that generates and analyzes the connectivity of the spatial superstructure is necessary.

5.2.4 *Feasible Paths*

A feasible path is a collection of nodes and edges that enables reaching final products from primary products. A feasible path can be derived by a reduction of nodes and edges of a max p-graph or superstructure by leveraging graph-theoretic concepts. For example, a feasible path is a subgraph $\mathcal{G}^f = (\mathcal{N}^f, \mathcal{E}^f)$ of the superstructure graph $\mathcal{G}^s$ (i.e., $\mathcal{G}^f \subseteq \mathcal{G}^s$) where $\mathcal{N}^f$ is a subset of $\mathcal{N}^s$ and $\mathcal{E}^f$ is a subset of $\mathcal{E}^s$. We present a feasible path obtained from an example max p-graph and superstructure in Figure 5.4. Compared to the number of possible paths derived from a max p-graph, the number of feasible paths from a superstructure is much larger because we now consider multiple copies of technologies (which enables more combinations). It is also important to highlight that, any feasible path in a superstructure (and spatial superstructure), has to be a feasible path for the max p-graph (because the superstructure inherits the product-tech connectivity). This observation is key in building superstructures that avoid spurious (infeasible) paths.

We can similarly derive any feasible paths from a spatial superstructure between primary products, techs, and final products as shown in Figure 5.5. Here, a couple of copies of the same technology $\mathcal{T}_{i_3,i_1,j_1}$ that produces product i_3 is located at location B (nodes

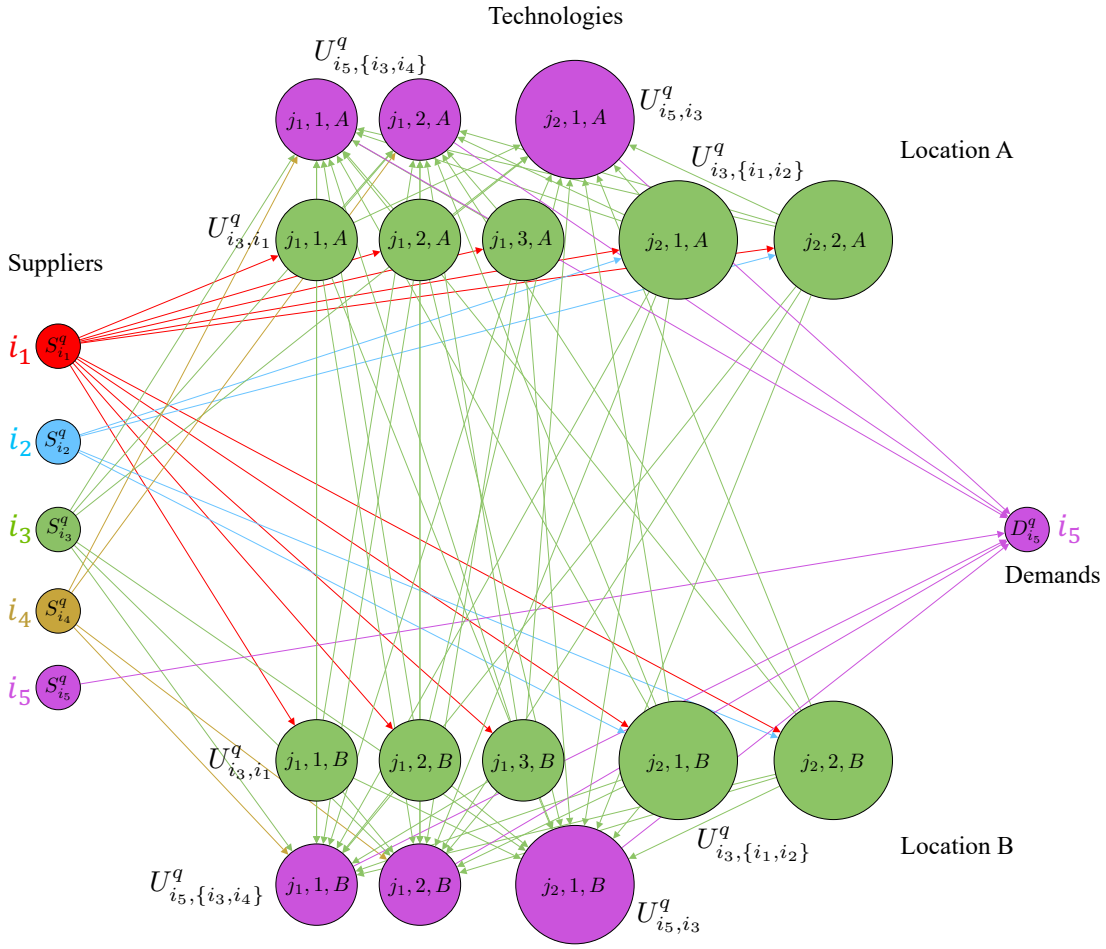


Figure 5.3: Illustration of a spatial superstructure (associated with max p-graph of Figure 5.1 and superstructure of Figure 5.2) showing dependencies between products and technologies across geographical locations.

$U^q_{i_3,i_1,j_1,1,B}$ and $U^q_{i_3,i_1,j_1,2,B}$), and two other technologies $\mathcal{T}_{i_3,\{i_1,i_2\},j_2}$ (node $U^q_{i_3,\{i_1,i_2\},j_2,1,A}$) and $\mathcal{T}_{i_5,\{i_3,i_4\},j_2}$ (node $U^q_{i_5,\{i_3,i_4\},j_2,1,A}$) that produces i_3 and final product i_5 are located at location B. As we add spatial information to the superstructure, combinations of techs across different locations are now possible and thus the number of possible feasible paths becomes even larger. Among all the feasible paths, an optimal design is a feasible path that takes into account tech and transport costs and modularity. As we consider complicated interdependencies between products and different capital and transportation cost due to potential locations, finding an optimal design is not immediately obvious from the spa-

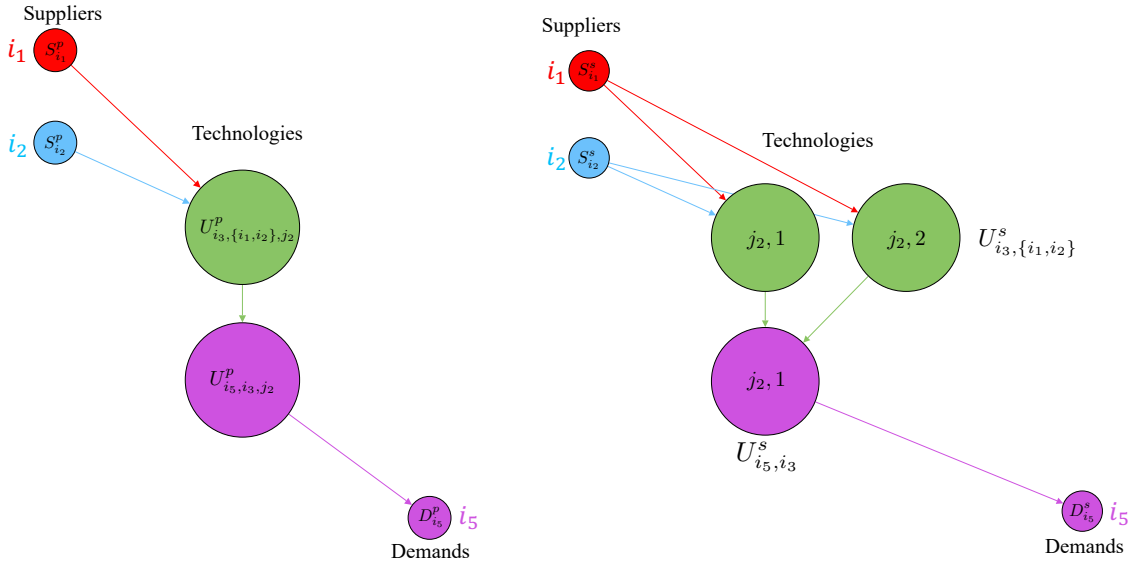


Figure 5.4: Example of a feasible path obtained from a max p-graph (left) and from a superstructure (right).

tial superstructure. Therefore, one needs to rely on optimization techniques to identify optimal paths, as we describe next.

5.3 Optimization Formulation for Finding Optimal System Designs

In this section, we derive an optimization formulation that aims to identify the hierarchy of products in the system. This step is essential for computing the amount of each product and the number of technologies required. This information is in turn required to generate superstructures and spatial superstructures. We then proceed to introduce an optimization formulation to obtain an optimal system design.

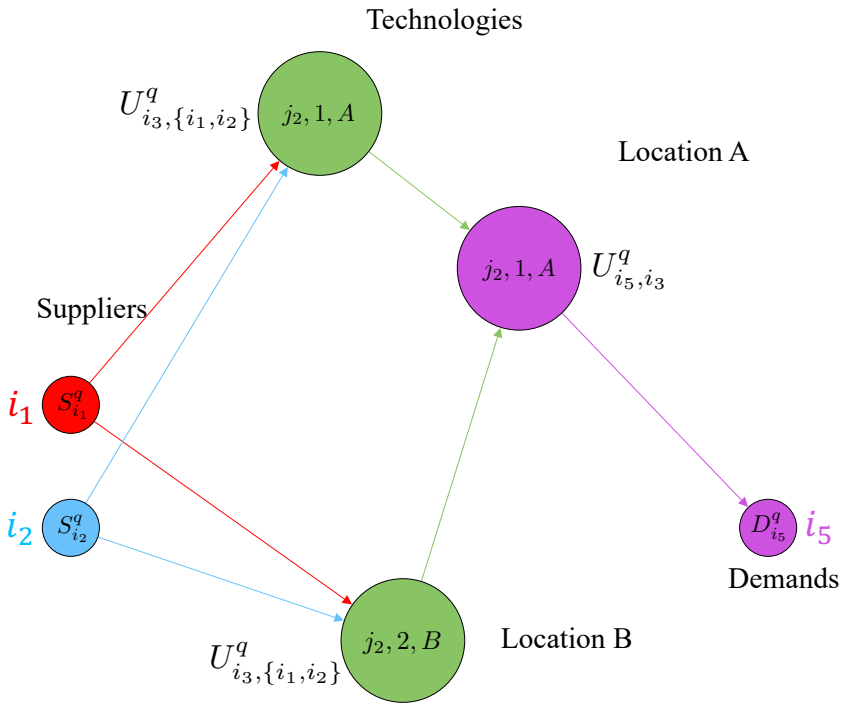


Figure 5.5: Example of a feasible path obtained from a spatial superstructure.

5.3.1 *Computing a Product Hierarchy*

Generating a superstructure and a spatial superstructure requires that we compute the number of all possible pathways based on their product-tech connectivity encoded in the max p-graph. Here, we assume that there are no cycles in the dependencies in the underlying max p-graph. For instance, if producing product i_1 requires some i_2, producing i_2 cannot require product i_1. With this, we can define a product hierarchy that moves from primary products to intermediate products and to final products. We use $\beta^{i'}_{i,t}, i' \in \mathcal{I}, i \in \Omega_t, t \in \mathcal{T}$ to represent that producing a unit of product $i \in \Omega_t$ requires $\beta^{i'}_{i,t}$ units of product i'. These quantities can be interpreted as technology yield/transformation factors.

Taking the example form the previous section, we consider the product-tech dependencies of a system that involves supplies of i_1 and i_2 and produces the intermediate product i_3 using techs $\mathcal{T}_{i_3,i_1,j_1}$ and $\mathcal{T}_{i_3,\{i_1,i_2\},j_2}$ and a system that involves supplies of i_3 and

Table 5.1: Example of product dependencies in technologies.

	i_3		i_5	
	$\mathcal{T}_{i_3,i_1,j_1}$	$\mathcal{T}_{i_3,\{i_1,i_2\},j_2}$	$\mathcal{T}_{i_5,\{i_3,i_4\},j_1}$	$\mathcal{T}_{i_5,i_3,j_2}$
i_1	2	1	0	0
i_2	0	1.5	0	0
i_3	0	0	0.8	1
i_4	0	0	1	0

i_4 and produces the final product i_5 using techs $\mathcal{T}_{i_5,\{i_3,i_4\},j_1}$ and $\mathcal{T}_{i_5,i_3,j_2}$. The product dependencies are shown in Table 5.1. Rows in this table represent input products while the columns represent output products. Specifically, producing i_3 requires 2 units of i_1 using $\mathcal{T}_{i_3,i_1,j_1}$ ($\beta^{i_1}_{i_3,\mathcal{T}_{i_3,i_1,j_1}} = 2$) or 1 unit of i_1 and 1.5 unit of i_2 using $\mathcal{T}_{i_3,\{i_1,i_2\},j_2}$ ($\beta^{i_1}_{i_3,\mathcal{T}_{i_3,\{i_1,i_2\},j_1}} = 1$, and $\beta^{i_2}_{i_3,\mathcal{T}_{i_3,\{i_1,i_2\},j_1}} = 1.5$), and it is represented in entries 2 and 1 in the first row (i_1 as the feed)and first two columns (i_3 as the product) in the table. Producing i_5 requires 0.8 unit of i_3 and 1 unit of i_4 using $\mathcal{T}_{i_5,\{i_3,i_4\},j_1}$ ($\beta^{i_3}_{i_5,\mathcal{T}_{i_5,\{i_3,i_4\},j_1}} = 0.8$ and $\beta^{i_4}_{i_5,\mathcal{T}_{i_5,\{i_3,i_4\},j_1}} = 1$) or 1 unit of i_3 using $\mathcal{T}_{i_5,i_3,j_2}$ ($\beta^{i_3}_{i_5,\mathcal{T}_{i_5,i_3,j_1}} = 1$), and it is represented in entries 0.8 and 1 in the third row (i_3 as the feed) and last columns (i_5 as the product) of the table. Here, i_1, i_2 and i_4 can be seen as a primary product, i_3 can be seen as an intermediate product and i_5 can be treated as a product. In this case, i_3 has a higher hierarchy than i_1 and i_2 because producing it depends on these products. Product i_5 has the highest hierarchy since no other products depend on it. Because there are no dependencies between products i_1, i_2, and i_4, the hierarchy among them can be arbitrary. Therefore, a possible hierarchy of these five products is $\{i_1 : 3, i_2 : 3, i_3 : 2, i_4 : 3, i_5 : 1\}$. Obtaining the product hierarchy allows us to estimate the amount of each product and the number of techs needed for the system to satisfy a set of demands; this information is necessary for generating the superstructure and spatial superstructure. Unfortunately, when the product-tech dependency becomes complicated, such product hierarchical order is not easy to observe; therefore, we formulate

an optimization problem that determines the hierarchical level of each product.

We define a positive integer variable $x_i, i \in \mathcal{I}$ that represents the hierarchy of each product i. The optimization formulation that computes the product hierarchy is:

$$\min_{x} \quad \sum_{i \in \mathcal{I}} x_i \tag{5.3.2a}$$

$$\text{s.t. } x_i \leq x_{i'} - 1, t \in \mathcal{T}, i \in \Omega_t, i' \in \mathcal{K}_t \tag{5.3.2b}$$

$$x_i \geq 1, i \in \mathcal{I}, \tag{5.3.2c}$$

Minimizing the objective (5.3.2a) ensures that the hierarchies for all products are consecutive numbers. Constraint (5.3.2b) indicates that, if producing product i requires product i', the hierarchy of i should be higher than the hierarchy of i'. The final constraint makes sure that the highest hierarchy starts with a value of one. This computation of the hierarchical level for each product aids the computation of the number of each techs possibly required, which is necessary information for generating the superstructures. Then, without further computations, we are able to derive a connectivity matrix (adjacency matrix) for all the nodes (products and techs) based on the information of products involved and their interdependencies, and thus build the graph representation of the superstructures.

We define the adjacency matrix for the superstructure and spatial superstructure as $\mu_{k,k'}, k, k' \in \mathcal{N}^s$ and $\mu_{k,k'}, k, k' \in \mathcal{N}^q$, respectively. The adjacency matrix is a fundamental quantity that encodes the topology of the superstructures.

5.3.2 *Computing Optimal Designs from Superstructures*

We first derive an optimization formulation to identify an optimal design from a superstructure that minimizes cost and maximizes modularity; this will generate a modular process design, as a superstructure does not encode locational context. We will then extend the formulation to identify an optimal design for a supply chain from a spatial superstructure.

Cost-Minimizing Optimal Design

A feasible path (a process design) is obtained by extracting a subgraph (nodes and edges) from the superstructure graph. Our goal is that this feasible path minimizes cost and maximizes modularity). The graph representation of the superstructure will allow us to derive an intuitive (and computable) measure for modularity, this measure will implicitly capture logistical flexibility (e.g., by capturing module sizes and connectivity).

We consider the overall system cost is the net present value for the annualized capital and operational cost. Associated with each tech $t \in \mathcal{T}$, we define the installation cost α_t^{ξ} and we define the capacity of tech t that produces product $i \in \Omega_t$ as $\xi_{t,i}$. For each node $u \in \mathcal{U}^s$ and its associated tech τ_u, we define the capital cost, operational cost, and capacity as $\alpha_{\tau_u}^{\xi}$, $\alpha_{\tau_u}^{o}$, and ξ_{τ_u,Ω_u} and we use the short-hand notation α_u^{ξ}, α_u^{o}, and $\xi_{u,i}, i \in \Omega_u$ (the capital cost is annualized with factor ϵ_u). The unit cost of each material is defined as $\alpha_i^{\rho}, i \in \mathcal{R}$ and the required amount of final product i is $\delta_i, i \in \mathcal{P}$. For simplicity, we assume that the cost for every connection/transport, denoted as α^f, is the same regardless of the product and scales linearly with the amount of product that it carries. The disposal cost of any excess product is denoted as $\alpha_i^d, i \in \mathcal{I}$; this disposal cost allows us to capture potential environmental impacts (e.g., carbon emissions).

We define a collection of continuous variables $f_{k,k'}, k, k' \in \mathcal{N}^s$ representing the flow of product from node k to node k'. We define a continuous variable $v_i, i \in \mathcal{R}$ that represents the amount of each primary product purchased from suppliers. We define a binary variable $y_u, u \in \mathcal{U}^s$ such that $y_u = 1$ if node u (a unit) is selected as part of the design, and $y_u = 0$ otherwise.

Under these definitions, the total annualized cost is:

$$\begin{aligned} C = & \sum_{u \in \mathcal{U}^s} (\epsilon_u \cdot \alpha_u^{\xi} + \alpha_u^{o}) \cdot y_u + \sum_{s \in \mathcal{S}^s, i \in \Omega_s} \alpha_i^{\rho} \cdot v_i + \sum_{k,k' \in \mathcal{N}^s} \alpha^f \cdot f_{k,k'} \\ & + \sum_{u \in \mathcal{U}^s, i \in \Omega_u} (\xi_u \cdot y_u - \sum_{k \in \mathcal{N}^s \text{ if } i \in \mathcal{K}_k} f_{u,k}) \cdot \alpha_i^d. \end{aligned} \tag{5.3.3}$$

The total cost captures installation cost, the cost of purchasing primary products, the cost of transport, and the cost of waste disposal. The installation cost implicitly captures economies of scale (as it captures technology cost based on size/capacity).

There are a couple of constraint sets that contribute to the formulation. The first set of constraints ensures that the feasible path is derived from the superstructure and can be expressed as:

$$f_{k,k'} \leq M \cdot \mu_{k,k'}, k, k' \in \mathcal{N}^s \tag{5.3.4}$$

where M is a sufficiently large coefficient; this constraint reduces the feasible region of the problem. The second set of constraints are the product balances at the graph nodes:

$$\sum_{k \in \mathcal{N}^s} f_{s,k} \leq v_i, s \in \mathcal{S}^s, i \in \Omega_s \tag{5.3.5a}$$

$$\sum_{k' \in \mathcal{N}^s} f_{u,k'} \leq \xi_{u,i} \cdot y_u, u \in \mathcal{U}^s, i \in \Omega_u \tag{5.3.5b}$$

$$\sum_{k \in \mathcal{N}^s} f_{k,u} = \xi_{u,i'} \cdot \beta^{i'}_{i,\tau_u} \cdot y_u, u \in \mathcal{U}^s, i' \in \mathcal{K}_u, i \in \Omega_u \tag{5.3.5c}$$

$$\sum_{k \in \mathcal{N}^s} f_{k,d} = \delta_{i'}, d \in \mathcal{D}^s, i' \in \mathcal{K}_d. \tag{5.3.5d}$$

Constraint (5.3.5a) is the product balance for supplier nodes, (5.3.5b) and (5.3.5c) ensure the inlet and outlet balance for tech nodes, and constraint (5.3.5c) is the balance for the demand node.

To illustrate the definition of the variables and constraints in the optimization formulation, we will use the same example discussed in the previous section. Consider that an optimal design is derived from a superstructure as shown in Figure 5.6. Nodes in the design on the right are marked from k_1 through k_5, where k_1 and k_2 are suppliers, k_3, k_4 and k_5 are technologies and k_6 is demand. If nodes k_3 and k_4 (copies of the same technology) take 400 units of primary product i_1 and 200 units of i_2 and produces 200 units of i_3. Product streams i_3 are then fed into node k_5 that produces 300 units of final product i_5 that is required by the market. The non-zero entries for each decision variables f, v, and

y are highlighted.

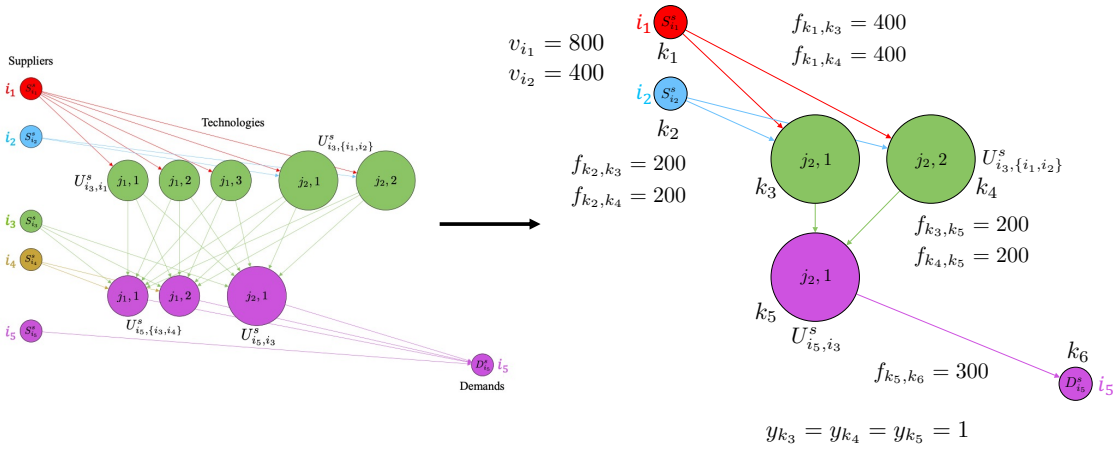


Figure 5.6: Illustration of notation for optimal design from superstructure.

With these definitions and constraints, we formulate the optimal design problem:

$$\min_{f,v,y} \; C \tag{5.3.6a}$$

$$\text{s.t. } f_{k,k'} \leq M \cdot \mu_{k,k'}, k, k' \in \mathcal{N}^s \tag{5.3.6b}$$

$$\sum_{k \in \mathcal{N}^s} f_{s,k} \leq v_i, s \in \mathcal{S}^s, i \in \Omega_s \tag{5.3.6c}$$

$$\sum_{k' \in \mathcal{N}^s \text{ if } i \in \mathcal{K}_{k'}} f_{u,k'} \leq \xi_{u,i} \cdot y_u, u \in \mathcal{U}^s, i \in \Omega_u \tag{5.3.6d}$$

$$\sum_{k \in \mathcal{N}^s} f_{k,u} = \xi_{u,i'} \cdot \beta^{i'}_{i,\theta_u} \cdot y_u, u \in \mathcal{U}^s, i' \in \mathcal{K}_u, i \in \Omega_u \tag{5.3.6e}$$

$$\sum_{k \in \mathcal{N}^s} f_{k,d} = \delta_{i'}, d \in \mathcal{D}^s, i' \in \mathcal{K}_d \tag{5.3.6f}$$

$$f_{k,k'} \geq 0, k, k' \in \mathcal{N}^s \tag{5.3.6g}$$

This formulation uses the superstructure connectivity to reduce the feasible space of the problem and is concise to read and easy to understand. We will now expand this formulation by incorporating a modularity measure as a trade-off of cost.

Cost-Minimizing Optimal Modular Design

The modularity measure that we adopt is a modified version of that presented in (Shao and Zavala, 2020). This measure is directly derived from graph-theoretical principles, can be computed using mixed-integer optimization techniques, and captures aspects of relevance in the context of manufacturing systems and supply chains. Specifically, the measure captures dimension (size) of modules, which enables capturing the fact that such modules should be transportable. The modularity measure proposed is computationally more suitable for system design.

The modularity measure is ccomputed using concept of graph coverage; this is done by formulating a mixed integer optimization problem that minimizes the number of inter-modular edges relative to the total number of edges (intra- and inter-modular). In other words, the measure aims to capture the ability of assembling/disassembling a system. The measure is defined as M_n with n being the predefined number of modules. The measure M_n has a range has a range of $[0, 1]$; $M_n = 1$ being the most modular system possible and $M_n = 0$ represents the least modular system possible.

In addition to the previous attributes of tech $t \in \mathcal{T}$, we define a dimension (physical size) of the tech represented as γ_t. Similarly, for each node $u \in \mathcal{U}^s$ and its associated tech $t = \tau_u$, the dimension is represented as γ_{τ_u}, abbreviated as γ_u. The set of modules is defined as $\mathcal{L} = \{1, 2, ..., n\}$ and $|\mathcal{L}| = n$ (the number of modules). We impose dimensionality constraints for each module by defining $\bar{D}$ and $\underline{D}$ as upper and lower limits.

To compute the *modularity measure* of a given feasible path (a potential design), we define the collection of binary parameters $a_{u,u'}, u, u' \in \mathcal{U}^s$, which represent the adjacency matrix of the subgraph associated to the feasible path. We define the collection of binary variables $a^m_{u,u',l}, u, u' \in \mathcal{U}^s, l \in \mathcal{L}$; here, $a^m_{u,u',l} = 1$ indicates that nodes u, u' are connected and they are both in module l. We define the binary variable $y_{u,l}, u \in \mathcal{U}^s, l \in \mathcal{L}$ such that $y_{u,l} = 1$ if node (unit) u exists in the design and appears in module l.

The measure that we propose is computed by minimizing the following function for

a pre-defined number of modules n:

$$M_n = \frac{\sum_{u,u' \in \mathcal{U}^s, l \in \mathcal{L}} a^m_{u,u',l}}{\sum_{u,u' \in \mathcal{U}^s} a_{u,u'}}. \tag{5.3.7}$$

We note that this modularity measure is different from that reported in (Shao and Zavala, 2020), which is:

$$M_n = \frac{\sum_{u,u' \in \mathcal{U}^s} \pi_{u,u'} \cdot a_{u,u'}}{\sum_{u,u' \in \mathcal{U}^s} a_{u,u'}}. \tag{5.3.8}$$

where π is the membership variable matrix. Here, we have that the entry $\pi_{u,u'}$ is 1 if node u and u' are in the same module. We highlight that, in the work of (Shao and Zavala, 2020), the adjacency $a_{u,u}$ is a fixed parameter and $\sum_{u,u' \in \mathcal{U}^s} a_{u,u'} = 2m$, where m is the number of edges in the graph. However, in the design context discussed here, the adjacency $a_{u,u}$ is a variable (affected by the design selection). As such, the second modularity measure would be computationally difficult to implement. This motivates our desire to use the first modularity measure (we will see that this is easier to implement).

We now proceed to show that the modularity measures are equivalent; we establish this result by showing that:

$$\sum_{l \in \mathcal{L}} a^m_{u,u',l} = \pi_{u,u'} \cdot a_{u,u'}. \tag{5.3.9}$$

In other words, we aim show that the numerators of both modularity measures are equivalent. Specifically, we would like to show that both numerators are binary and that, when the numerator on the left takes a value of 1, the numerator on the right also takes a value of 1.

The right numerator is binary because both terms are binary and thus their product is binary. For the left numerator we have that, by definition, there are no overlapping

modules (each node can exist only in one module). For a module $l' \in \mathcal{L}$ we thus have that:

$$a^m_{u,u',l'} = 1 \iff \sum_{l \in \mathcal{L}} a^m_{u,u',l} = 1 \tag{5.3.10}$$

and

$$\sum_{l \in \mathcal{L}} a^m_{u,u',l} \leq 1. \tag{5.3.11}$$

Because the variable a^m is also binary, we have that it can only take values of 0 and 1. We then have that, by definition, the term $a^m_{u,u',l'}$ takes a value of 1 if and only if node u connects to node u' and they are both in module l. Therefore, for a module $l' \in \mathcal{L}$, we have:

$$a^m_{u,u',l'} = 1 \iff a_{u,u'} = 1 \text{ and } \pi_{u,u'} = 1. \tag{5.3.12}$$

Combining these expressions:

$$\sum_{l \in \mathcal{L}} a^m_{u,u',l} = 1 \iff a_{u,u'} = 1 \text{ and } \pi_{u,u'} = 1. \tag{5.3.13}$$

Therefore, we can rewrite the above expression as:

$$\sum_{l \in \mathcal{L}} a^m_{u,u',l} = a_{u,u'} \cdot \pi_{u,u'} \tag{5.3.14}$$

which establishes the equivalence.

Note that the new modularity measure proposed is nonlinear (fractional) and would make the design formulation intractable if this is added directly as an objective function. Interestingly, however, when the measure is used as a constraint, this can be reformulated

in linear form as:

$$\sum_{u,u' \in \mathcal{U}^s, l \in \mathcal{L}} a^m_{u,u',l} \geq \epsilon_M \cdot \sum_{u,u' \in \mathcal{U}^s} a_{u,u'}, \tag{5.3.15}$$

where ϵ_M is a desired threshold value for the modularity measure. This observation is relevant because the design formulation is multi-objective (minimize cost and maximize modularity). Using an ϵ-constrained method to compute the Pareto frontier thus provides a natural approach to deal with the modularity measure.

In addition to the modularity constraint and the constraints in formulation (5.3.6), new constraints are added the impact of the assignment of nodes into modules on process variables. The logic between continuous flow variables and the binary adjacency variables is:

$$f_{k,k'} \leq M \cdot a_{k,k'}, k, k' \in \mathcal{N}^s \tag{5.3.16}$$

$$a_{k,k'} \leq f_{k,k'}, k, k' \in \mathcal{N}^s. \tag{5.3.17}$$

The logic for variable y (a node can only exist in one module), variable a^m, and a (a node u connects to any node u' and they all belong to a module l only if node u and u' both exist) are:

$$\sum_{l \in \mathcal{L}} y_{u,l} \leq 1, u \in \mathcal{U}^s \tag{5.3.18}$$

$$\sum_{u' \in \mathcal{U}^s} a^m_{u,u',l} + \sum_{u' \in \mathcal{U}^s} a^m_{u',u,l} \leq M \cdot y_{u,l}, u \in \mathcal{U}^s, l \in \mathcal{L} \tag{5.3.19}$$

$$\sum_{l \in \mathcal{L}} a^m_{u,u',l} \leq a_{u,u'}, u, u' \in \mathcal{U}^s \tag{5.3.20}$$

The constraint that governs the upper and lower bound for the dimensionality of each module is:

$$\underline{D} \leq \sum_{u \in \mathcal{U}^s} y_{u,l} \cdot \gamma_u \leq \bar{D}, l \in \mathcal{L}. \tag{5.3.21}$$

Figure 5.7 presents an example to demonstrate the definitions of variables and constraints. Here, we show a modular division of the optimal design where k_3 and k_5 are in module l_1 and k_4 is in module l_2. It this worth noticing is that only one entry of variable a^m is 1 since only the edge that connects node k_3 and k_5 is counted as the edge within modules.

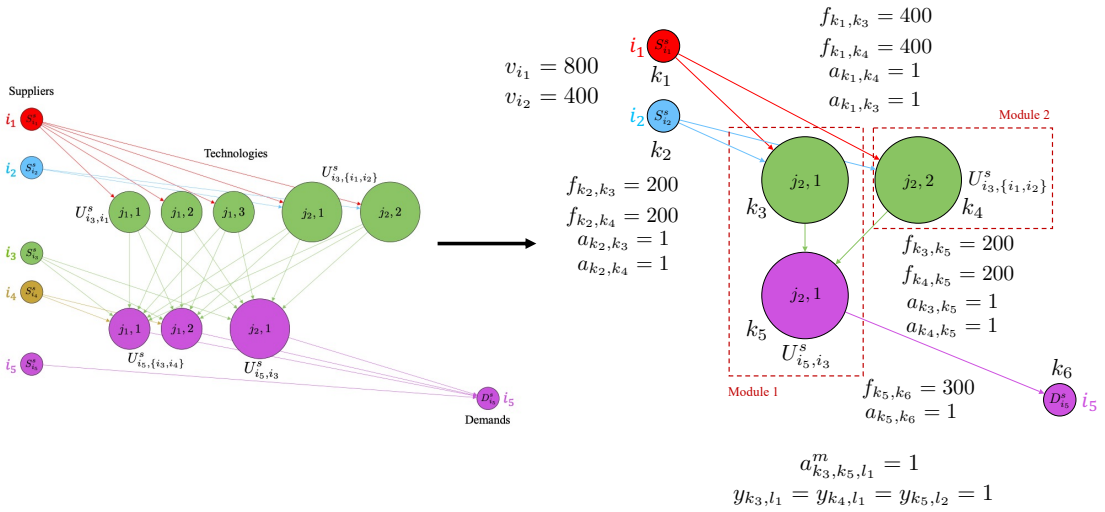


Figure 5.7: Illustration of notation for optimal design (left) obtained from superstructure (right).

The optimization formulation to select the cost-optimal design given a modularity threshold ϵ_M is:

$$\min_{f,v,a,a^m,y} \quad C \tag{5.3.22a}$$

$$\text{s.t. } f_{k,k'} \leq M \cdot \mu_{k,k'}, k, k' \in \mathcal{N}^s \tag{5.3.22b}$$

$$\sum_{k \in \mathcal{N}^s} f_{s,k} \leq v_i, s \in \mathcal{S}^s, i \in \Omega_s \tag{5.3.22c}$$

$$\sum_{k' \in \mathcal{N}^s} f_{u,k'} \leq \xi_{u,i} \cdot \sum_{l \in \mathcal{L}} y_{u,l}, u \in \mathcal{U}^s, i \in \Omega_u \tag{5.3.22d}$$

$$\sum_{k \in \mathcal{N}^s} f_{k,u} = \xi_{u,i'} \cdot \beta^{i'}_{i,\tau_u} \cdot \sum_{l \in \mathcal{L}} y_{u,l}, u \in \mathcal{U}^s, i' \in \mathcal{K}_u, i \in \Omega_u \tag{5.3.22e}$$

$$\sum_{k \in \mathcal{N}^s} f_{k,d} = \delta_{i'}, d \in \mathcal{D}^s, i' \in \mathcal{K}_d \tag{5.3.22f}$$

$$f_{k,k'} \leq M \cdot a_{k,k'}, k, k' \in \mathcal{N}^s \tag{5.3.22g}$$

$$a_{k,k'} \leq f_{k,k'}, k, k' \in \mathcal{N}^s \tag{5.3.22h}$$

$$\sum_{l \in \mathcal{L}} y_{u,l} \leq 1, u \in \mathcal{U}^s \tag{5.3.22i}$$

$$\sum_{u' \in \mathcal{U}^s} a^m_{u,u',l} + \sum_{u' \in \mathcal{U}^s} a^m_{u',u,l} \leq M \cdot y_{u,l}, u \in \mathcal{U}^s, l \in \mathcal{L} \tag{5.3.22j}$$

$$\sum_{l \in \mathcal{L}} a^m_{u,u',l} \leq a_{u,u'}, u, u' \in \mathcal{U}^s \tag{5.3.22k}$$

$$\underline{D} \leq \sum_{u \in \mathcal{U}^s} y_{u,l} \cdot \gamma_u \leq \bar{D}, l \in \mathcal{L} \tag{5.3.22l}$$

$$\sum_{u,u' \in \mathcal{U}^s, l \in \mathcal{L}} a^m_{u,u',l} \geq \epsilon_M \cdot \sum_{u,u' \in \mathcal{U}^s} a_{u,u'} \tag{5.3.22m}$$

$$f_{k,k'} \geq 0, k, k' \in \mathcal{N}^s \tag{5.3.22n}$$

Incorporating the modularity measure in the design formulation increases the number of variables and constraints, which are required for the calculation of intra-modular edges. However, this formulation gives us interesting insights on connectivity, decentralization, and transportability of the process design. Formulation (5.3.6) and (5.3.22) find the optimal modular design based on the superstructure graph (that does not account for spatial information). We will now introduce formulations that find the optimal feasible path with spatial information, and we would like to show that with the concept of spatial superstructure, these formulations are similar to formulations (5.3.6) and (5.3.22).

5.3.3 *Computing Optimal Designs from Spatial Superstructures*

We then consider the situation that we not only assign technologies to modules, but also put them at different locations. Here, we consider two potential factors that might change the geological preference for different technologies. First, the installation cost may be different for the same technology at different locations due to the different cost of land. Second, the transportation cost from one location to another may be different for the same product. Specifically, if technologies connected to each other are installed at the same location, the short-range transport of product can be achieved using pipelines or other

simple methods. If they are placed at different locations, trucks or trains may be utilized to long-range transport of the product. Alternative for short- and long-range transport options are directly captured by our formulation. Also, the primary products are usually supplied at some certain locations and the final products are usually transported to the location of demand. We further assume that technologies within the same module should be placed together in the same location and multiple modules can be placed at the same location. We first formulate the problem that solves for the cost-minimizing optimal spatial feasible path and then incorporate the modularity measure for an optimal modular design.

Cost-Minimizing Optimal Design with Spatial Information

The graph-theoretic representation of the spatial superstructure makes the notation of the optimization formulation directly analogous to that of the superstructure. As such, we briefly discuss all the necessary definitions for the formulation.

The capacity associated with technology $t = \tau_u$ for node $u \in \mathcal{U}^q$ are abbreviated as ξ_u. The installation cost associated with each technology $t \in \mathcal{T}$ at location $g \in \mathcal{G}_t$ is defined as $\alpha^{\xi}_{t,g}$. For the node $u \in \mathcal{U}^q$ associated with techτ_u, the installation cost can be represented as $\alpha^{\xi}_{\tau_u,\phi_u}$, abbreviated similarly as α^{ξ}_{u}. The unit cost of each product is α^{ρ}_{i} and the demand of the market for each product $i \in \mathcal{P}$ is defined as δ_i, and the cost of disposal for each product is defined as α^{d}_{i}. We redefine the cost for connection/flow of product i from location g to location g' which stands for the cost of transportation as $\alpha^{f}_{i,g,g'}, i \in \mathcal{I}, g, g' \in \mathcal{G}$. Note that if $g = g'$, we obtain the short-term transport cost (as in a typical process).

The variables defined for the spatial superstructure problem have a similar interpretation as those of the superstructure but we can now attribute locational context. Specifically, we define a continuous variable $f_{k,k'}, k, k' \in \mathcal{N}^q$ representing the flow of product from node k at location ϕ_k to node k' at location $\phi_{k'}$. We also define a continuous variable for the purchase of each primary product as $v_i, i \in \mathcal{R}$. We define a binary variable matrix

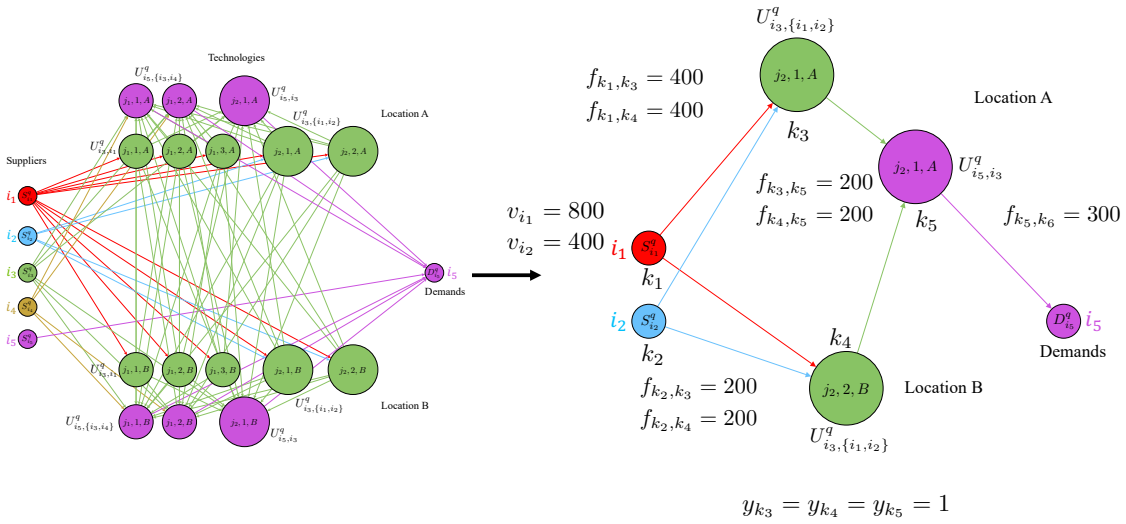


Figure 5.8: Illustration of notation for optimal design (right) obtained from spatial superstructure (left).

$y_u, u \in \mathcal{U}^q$ such that $y_u = 1$ if node u is selected as part of the design located at location ϕ_u, and $y_u = 0$ otherwise. Therefore, the total annualized cost of the system can be written as:

$$\begin{aligned} C = & \sum_{u \in \mathcal{U}^q} (\epsilon \cdot \alpha^{\xi}{}_u + \alpha^o_u) \cdot y_u + \sum_{s \in \mathcal{S}^q, i \in \Omega_s} \alpha^p_i \cdot v_i + \sum_{i \in \Omega_k, k \in \mathcal{N}^q, k' \in \mathcal{N}^q} \alpha^f_{i,\phi_k,\phi_{k'}} \cdot f_{k,k'} \\ & + \sum_{u \in \mathcal{U}^q, i \in \Omega_u} (\xi_u \cdot y_u - \sum_{k \in \mathcal{N}^q \text{ if } k \in \mathcal{K}_k} f_{u,k}) \cdot \alpha^d_i. \end{aligned} \tag{5.3.23}$$

The only difference in this cost function is that we can now capture transport cost. An illustration of the problem variables is provided in Figure 5.8.

The formulation to obtain a cost-minimizing design from the spatial superstructure is:

$$\min_{f,v,y} \; C \tag{5.3.24a}$$

$$\text{s.t. } f_{k,k'} \leq M \cdot \mu_{k,k'}, k, k' \in \mathcal{N}^q \tag{5.3.24b}$$

$$\sum_{k \in \mathcal{N}^q} f_{s,k} \leq v_i, s \in \mathcal{S}^q, i \in \Omega_s \tag{5.3.24c}$$

$$\sum_{k' \in \mathcal{N}^q \text{ if } i \in \mathcal{K}_{k'}} f_{u,k'} \leq \xi_{u,i} \cdot y_u, u \in \mathcal{U}^q, i \in \Omega_u \tag{5.3.24d}$$

$$\sum_{k \in \mathcal{N}^q} f_{k,u} = \xi_{u,i'} \cdot \beta_{i,\tau_u}^{i'} \cdot y_u, u \in \mathcal{U}^q, i' \in \mathcal{K}_u, i \in \Omega_u \tag{5.3.24e}$$

$$\sum_{k \in \mathcal{N}^q} f_{k,d} = \delta_{i'}, d \in \mathcal{D}^q, i' \in \mathcal{K}_d \tag{5.3.24f}$$

$$f_{k,k'} \geq 0, k, k' \in \mathcal{N}^q \tag{5.3.24g}$$

The spatial superstructure is encoded $\mu_{k,k'}, k, k' \in \mathcal{N}^q$. This optimization formulation is directly analogous to that in (5.3.6) but can be computationally more challenging to solve because one can account for multiple possible locations for techs, suppliers, and demands.

Cost-Minimizing Optimal Modular Design with Spatial Information

The modularity measure is directly analogous to the one defined previously but we need to specify additional information to account for location of modules. The dimension associated with technology $t = \tau_u$ for node $u \in \mathcal{U}^q$ are abbreviated as γ_u. We use set $\mathcal{L} = \{1, 2, ..., n\}$ for the set of modules, and $\bar{D}$ and $\underline{D}$ to represent the upper and lower limits of the dimensionality requirements. We use the binary variable matrix $a_{k,k'}, k, k' \in \mathcal{N}^q$ to represent the adjacency matrix of the feasible path (subgraph of the spatial superstructure). The binary variable $a^m_{u,u',l}, u, u' \in \mathcal{U}^q, l \in \mathcal{L}$ represents the relationship between node u, u' and module l. We define the binary variable $y_{u,l}, u \in \mathcal{U}^q, l \in \mathcal{L}$ such that $y_{u,l} = 1$ if node u belongs to module l at location ϕ_u and $y_{u,l} = 0$ otherwise. Finally, we define the binary variable collection $z_{l,g}, l \in \mathcal{L}, g \in \mathcal{G}_t$ such that if module l is placed at location g, $z_{l,g} = 1$ and $z_{l,g} = 0$ otherwise.

The modularity measure with predefined number of modules n modules is:

$$M_n = \frac{\sum_{u,u' \in \mathcal{U}^q, l \in \mathcal{L}} a^m_{u,u',l}}{\sum_{u,u' \in \mathcal{U}^q} a_{u,u'}}. \tag{5.3.25}$$

A couple of additional constrains (compared to formulation (5.3.22)) are added due to

the newly defined variable z. The first ensures that one module can only be placed at one location, and the second one ensures the logic between variable y and z (a node u that exists in module l only if the module l exists at location of node u), and they are expressed as:

$$\sum_{g \in \mathcal{G}_t} z_{l,g} \leq 1, l \in \mathcal{L} \tag{5.3.26}$$

$$y_{u,l} \leq z_{l,\phi_u}, u \in \mathcal{U}^q, l \in \mathcal{L}. \tag{5.3.27}$$

An illustration of the defined variables is shown in Figure 5.9; note that a couple of entries for the newly defined variable z are 0, as module l_1 exists at location A and module l_2 exists at location B.

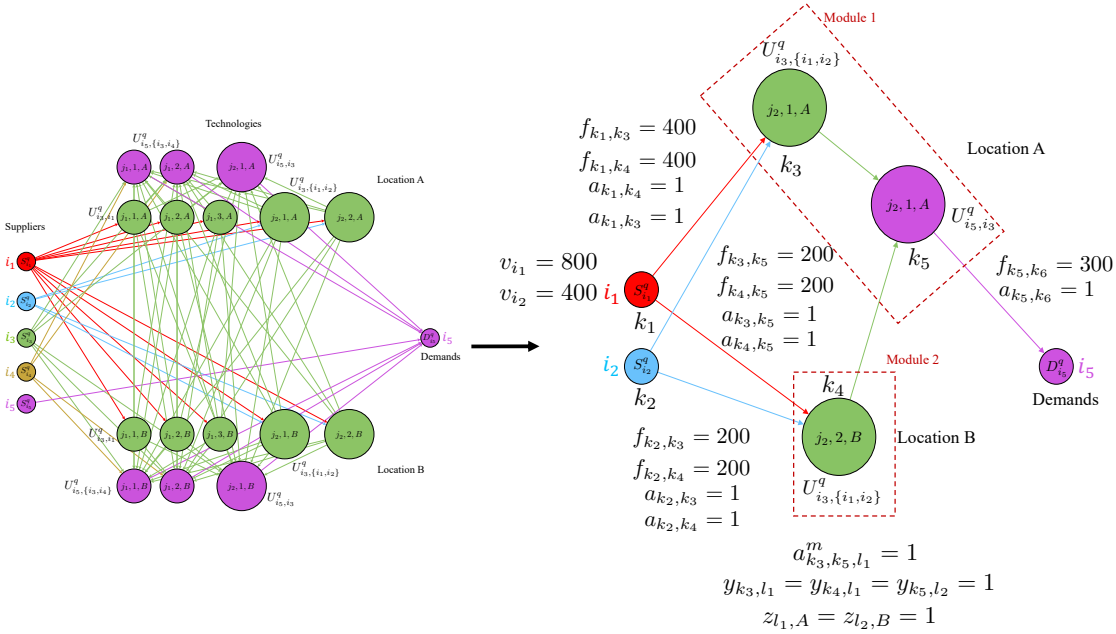


Figure 5.9: Illustration of notation for optimal modular design from spatial superstructure

With the previous definitions, the optimization formulation is:

$$\min_{f,v,a,a^m,y,z} \quad C \tag{5.3.28a}$$

$$\text{s.t. } f_{k,k'} \leq M \cdot \mu_{k,k'}, k, k' \in \mathcal{N}^q \tag{5.3.28b}$$

$$\sum_{k \in \mathcal{N}^q} f_{s,k} \leq v_i, s \in \mathcal{S}^q, i \in \Omega_s \tag{5.3.28c}$$

$$\sum_{k' \in \mathcal{N}^q} f_{u,k'} \leq \xi_{u,i} \cdot \sum_{l \in \mathcal{L}} y_{u,l}, u \in \mathcal{U}^q, i \in \Omega_u \tag{5.3.28d}$$

$$\sum_{k \in \mathcal{N}^q} f_{k,u} = \xi_{u,i'} \cdot \beta_{i,\tau_u}^{i'} \cdot \sum_{l \in \mathcal{L}} y_{u,l}, u \in \mathcal{U}^q, i' \in \mathcal{K}_u, i \in \Omega_u \tag{5.3.28e}$$

$$\sum_{k \in \mathcal{N}^q} f_{k,d} = \delta_{i'}, d \in \mathcal{D}^q, i' \in \mathcal{K}_d \tag{5.3.28f}$$

$$f_{k,k'} \leq M \cdot a_{k,k'}, k, k' \in \mathcal{N}^q \tag{5.3.28g}$$

$$a_{k,k'} \leq f_{k,k'}, k, k' \in \mathcal{N}^q \tag{5.3.28h}$$

$$\sum_{l \in \mathcal{L}} y_{u,l} \leq 1, u \in \mathcal{U}^q \tag{5.3.28i}$$

$$\sum_{u' \in \mathcal{U}^q} a_{u,u',l}^m + \sum_{u' \in \mathcal{U}^q} a_{u',u,l}^m \leq M \cdot y_{u,l}, u \in \mathcal{U}^q, l \in \mathcal{L} \tag{5.3.28j}$$

$$\sum_{l \in \mathcal{L}} a_{u,u',l}^m \leq a_{u,u'}, u, u' \in \mathcal{U}^q \tag{5.3.28k}$$

$$\underline{D} \leq \sum_{u \in \mathcal{U}^q} y_{u,l} \cdot \gamma_u \leq \bar{D}, l \in \mathcal{L} \tag{5.3.28l}$$

$$\sum_{u,u' \in \mathcal{U}^q, l \in \mathcal{L}} a_{u,u',l}^m \geq \epsilon \cdot \sum_{u,u' \in \mathcal{U}^q} a_{u,u'} \tag{5.3.28m}$$

$$f_{k,k'} \geq 0, k, k' \in \mathcal{N}^q \tag{5.3.28n}$$

$$\sum_{g \in \mathcal{G}_t} z_{l,g} \leq 1, l \in \mathcal{L} \tag{5.3.28o}$$

$$y_{u,l} \leq z_{l,\phi_u}, u \in \mathcal{U}^q, l \in \mathcal{L} \tag{5.3.28p}$$

This formulation is more comprehensive that the one based on a superstructure in that it delivers an optimal system design that not only captures techs needed but also their geographical location. In other words, this formulation simultaneously designs a supply chain and associated processes.

5.4 Case Study

We present a case study to illustrate how our optimization formulations can help automate the generation of superstructures and spatial superstructures and to identify optimal system designs with desired modularity. The study tries to identify an optimal supply chain design for plastic waste upcycling that takes municipal solid waste (MSW) as the input and produces ethylene, propylene and hydrogen as final products. The MILPs were solved using Gurobi (version 9.0.3) and were implemented in the Julia-based JuMP modeling framework. We use Gephi for graph manipulation and visualization. All optimization formulations are solved using a commercial laptop and the solving time is referred to the wall clock time. All code needed to reproduce the results can be found in https://github.com/zavalab/JuliaBox/tree/master/ModularDesign.

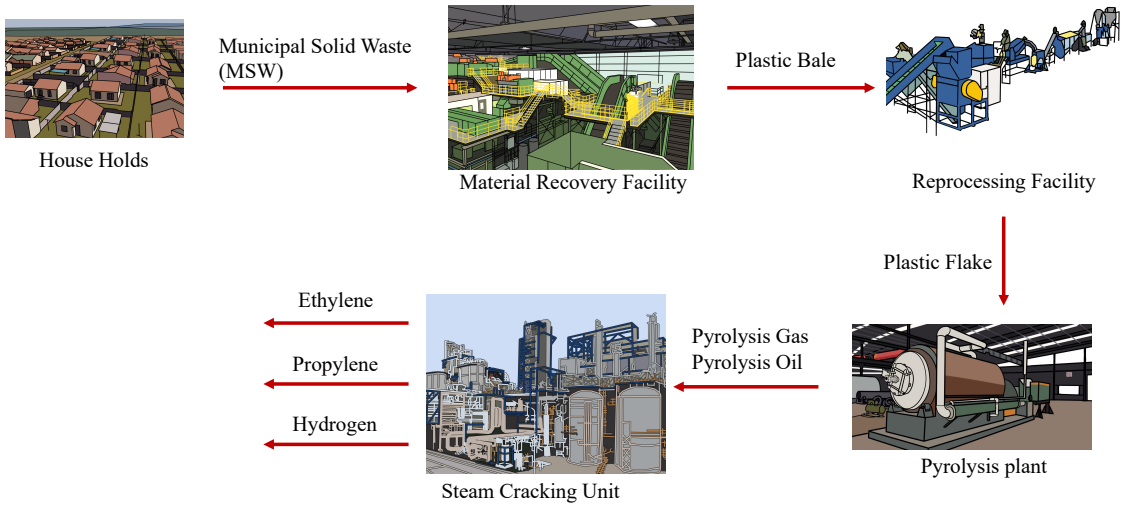


Figure 5.10: High-level view of processing tasks involved in plastic waste upcycling.

5.4.1 *Problem Setup and Material Hierarchy of the Process*

A high-level view of the processing tasks involved in plastic waste upcycling is provided in Figure 5.10. Here, MSW (denoted as i_1) collected from households is fed into a product

recovery facility (MRF) that obtains a plastic bale (i_2), the plastic bale goes through a reprocessing facility (RF) that cleans the bale and converts it into plastic flakes (i_3), a pyrolysis process (PP) takes the plastic flakes and converts these into pyrolysis gas (i_4) and pyrolysis oil (i_5), a steam cracking (SC) process obtains the final products, given by ethylene (i_6), propylene (i_7), and hydrogen (i_8). For this system, techs that produce the same products have the same interdependencies between products and they only differ in their capacities. Therefore, we eliminate the attribute of techs and the product dependencies between the different techs is shown in Table 5.2. Producing 1 unit of i_2 requires 7.69 unit of i_1; producing 1 unit of i_3 requires 1 unit of i_2; producing 1 unit of i_5 requires 1.29 unit of i_3; finally, 1 unit of i_6, i_7 or i_8 requires 3.81, 6.16 or 125 unit of i_5 respectively. Note that intermediate product i_4 is not used in the following process and is therefore considered a waste. Additional information is summarized in Table 5.3.

Table 5.2: Product interdependencies between technologies.

	i_2	i_3	i_5	i_6	i_7	i_8
i_1	7.69	0	0	0	0	0
i_2	0	1	0	0	0	0
i_3	0	0	1.29	0	0	0
i_5	0	0	0	3.81	6.16	125

Each installation cost, operating cost tech dimension and tech capacity is associated with each technology in the above row respectively. For example, for technology $T_{\{i_4,i_5\},i_3,j_1}$, its installation cost is \$ 0.46×10^8, operating cost is \$ 14 per unit of input i_3, dimension is 2, and capacities for producing product i_4 and i_5 are 169,000 tons and 39,000 tons, respectively. We assume that the operating cost for techs producing the same products but of different sizes/kinds is the same. It is worth emphasizing that, to capture economies of scale, the estimation of the installation cost is based on the so-called "2/3 rule" that is prevalent in cost estimation. Specifically, the rule that applies to the different technologies producing product i_2 can be expressed as:

Table 5.3: Data for plastic upcycling system.

Parameters	Values
Supplied Materials,$\mathcal{R}$	$[i_1, i_2, i_3, i_5]$
Products, $\mathcal{P}$	$[i_2, i_3, i_4, i_5, i_6, i_7, i_8]$
All Materials, $\mathcal{I}$	$[i_1, i_2, i_3, i_4, i_5, i_6, i_2, i_8]$
Technologies , $\mathcal{T}$	$[T_{i_2,i_1,j_1}, T_{i_2,i_1,j_2}, T_{i_2,i_1,j_3}]$ $[T_{i_3,i_2,j_1}, T_{i_3,i_2,j_2}, T_{i_3,i_2,j_3}]$ $[T_{\{i_4,i_5\},i_3,j_1}, T_{\{i_4,i_5\},i_3,j_2}, T_{\{i_4,i_5\},i_3,j_3}]$ $[T_{\{i_6,i_7,i_8\},i_5,j_1}, T_{\{i_6,i_7,i_8\},i_5,j_2}, T_{\{i_6,i_7,i_8\},i_5,j_3}]$
Installation Cost, α^{ξ} ($\times 10^8$ \$)	[0.27, 0.46, 0.70] [0.13, 0.30, 0.56] [0.46, 0.79, 1.20] [6.05, 9.17 13.90]
Operating Cost, α^{u} (\$ / unit of input)	8.87 44.19 14 71.8
Technology Dimension, γ	[3, 5, 8] [3, 5, 8] [2, 4, 8] [3, 5, 8]
Technology Capacity, ξ ($\times 10^4$ tons)	[24.2, 60.5, 120.9] [20.8, 52, 104] [[16.9, 3.9], [42.3, 9.8], [112.8, 26.3]] [[13.1, 8.1, 0.4], [26.1, 16.2, 0.80], [52.3, 32.3, 1.6]]
Purchasing Unit Cost, $\alpha_i^{\rho}, i \in \mathcal{R}$, (\$/unit)	[0, 250, 1300, 1100]
Disposal Cost, $\alpha_i^{d}, i \in \mathcal{I}$, (\$/ton)	[50, 40, 40, 400, 800, 0, 0, 0]
Required Production, $\delta_i, i \in \mathcal{I}$, (tons)	[0, 0, 0, 0, 0, 150000, 100000, 5000]
Project Duration, t_p (yrs)	20
Discount Rate, r	0.06
Annualization Factor, ϵ_u	0.087
Module Dimension Limits, $[\underline{D}, \bar{D}]$	[2, 12]

$$\left(\frac{\xi_{T_{i_2,i_1,j_1}}}{\xi_{T_{i_2,i_1,j_2}}}\right) = \left(\frac{\alpha^{\xi}_{T_{i_2,i_1,j_1}}}{\alpha^{\xi}_{T_{i_2,i_1,j_2}}}\right)^{\frac{3}{2}}. \tag{5.4.29}$$

In order to generate the superstructures of the system, we first use the formulation (5.3.2) to solve for the hierarchy of all products based on their interdependencies. The result shows that i_4, i_6, i_7, and i_8 have the highest hierarchy followed by i_5, i_3, i_2 and i_1 respectively. This makes sense because products i_6, i_7, and i_8 are the final product of the process, and for intermediate product i_4, even though it is the product of an intermediate process, it is the waste of the process that no other products depend on. Therefore, it has the highest hierarchy but it is not a required product in the process. Then, with the hierarchy of each product, we can compute the total amount of each product needed for the system and the number of all possible techs needed for the process. The information is summarized in Table 5.4.

With above information, we are now ready to generate the superstructures of the system and then identify an optimal feasible path. We first consider the design using superstructure without spatial information and solve the problem using formulation (5.3.6) and (5.3.22). Then, we solve for the spatial optimal design using formulation (5.3.24) and (5.3.28).

5.4.2 *Optimal System Design without Spatial Information*

The superstructure of this system is shown in Figure 5.11. Looking from top to bottom, we have the product hierarchy of the system, which takes MSW as the input, generates plastic bale/flakes and pyrolysis gas/oil as intermediate products, and generates ethylene, propylene and hydrogen as final outputs. Looking from left to right, we have four nodes representing the supplies of the four primary products, and in the middle nodes with different sizes represent techs of different capacities. The number of copies of each

Table 5.4: Results for hierarchy of products, quantity of products, and number of technologies.

Materials	Hierarchy	Required Amount (tons)	Number of Technologies
i_1	5	6.22×10^6	–
i_2	4	8.09×10^5	T_{i_2,i_1,j_1}: 4 T_{i_2,i_1j_2}: 2 T_{i_2,i_1,j_3}: 1
i_3	3	8.09×10^5	T_{i_3,i_2,j_1}: 4 T_{i_3,i_2,j_2}: 2 T_{i_3,i_2,j_3}: 1
i_5	2	6.25×10^5	$T_{\{i_4,i_5\},i_3,j_1}$: 4 $T_{\{i_4,i_5\},i_3,j_2}$: 2 $T_{\{i_4,i_5\},i_3,j_3}$: 1
i_4	1	0	/
i_6	1	1.5×10^5	$T_{\{i_6,i_7,i_8\},i_5,j_1}$: 2
i_7	1	1×10^5	$T_{\{i_6,i_7,i_8\},i_5,j_2}$: 1
i_8	1	5×10^3	$T_{\{i_6,i_7,i_8\},i_5,j_3}$: 1

techs coincides with the number shown in Table 5.4. Finally, the three nodes on the right represent the demands. As expected, the superstructure is dense due to the large number of possible techs.

We assume that the connectivity cost α^f is \$ 0.01 per unit of product that an edge carries. First, we used formulation (5.3.6) to solve for the cost-minimizing optimal supply chain design and the result is shown in Figure 5.12 . The design problem contains 1028 continuous variables and 25 binary variables, and contains 1111 constraints. This problem takes less than 0.01 second to solve. The result shows that the optimal design design contains 7 tech units and achieves an annualized cost of \$ 6.56×10^8. We can see that it chooses the tech with largest capacity for processes MRF, RF and PP as they are the most cost efficient units to satisfy the required amount of products.

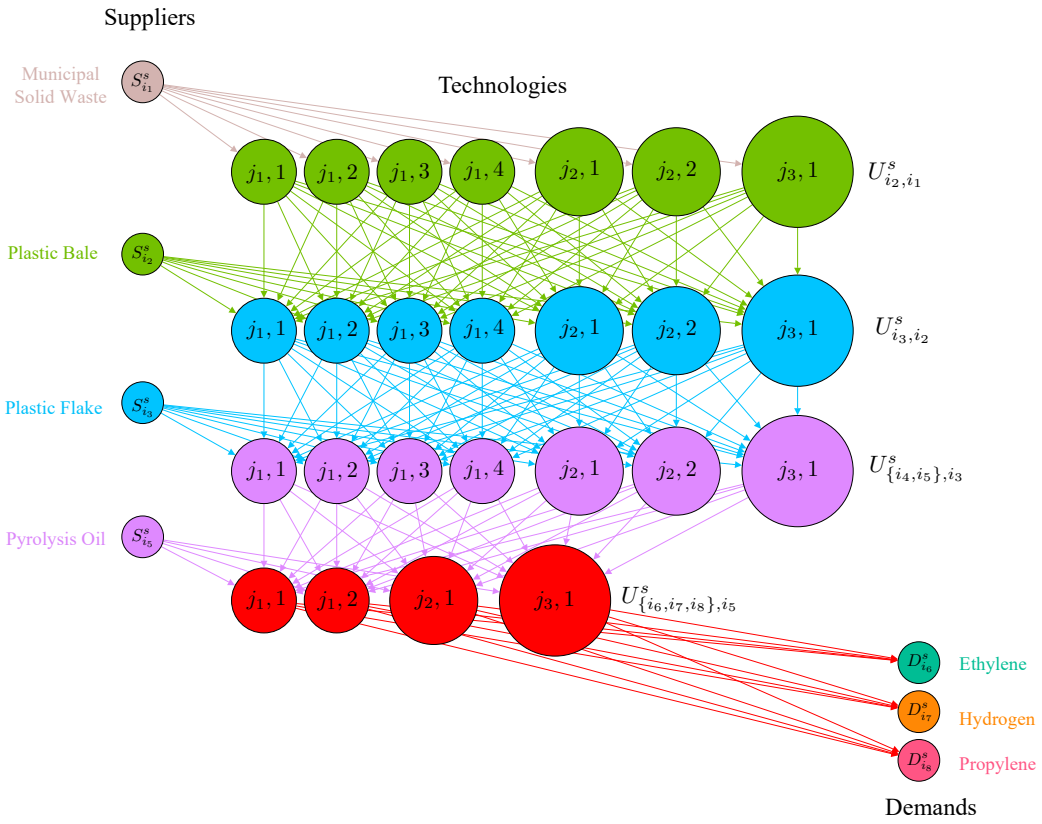


Figure 5.11: Superstructure for plastic upcycling system (no spatial information).

We then use optimization formulation (5.3.22) to solve for the cost-minimizing design that also achieves a certain degree of modularity. This problem has 1028 continuous variables, 3624 integer variables, and 3918 constraints. A couple of designs that correspond to different levels of modularity are shown in Figure 5.13. Module division in both cases are grouped by red dashed rectangles and nodes within 4 modules for each case are summarized in Table 5.5. Note that the optimal design shown on the left in Figure 5.13 contains identical techs as in Figure 5.12. This means that the cost-minimizing design without considering modularity achieves a modularity measure of 0.3. This is the maximum level of modularity that a valid system design can achieve, while achieving the minimum level of the cost (and we see that this value is quite low). We can also see that, as we require a higher degree of modularity, medium- and small-sized technologies start to be considered in the design. We summarize the results for different combinations of modularity

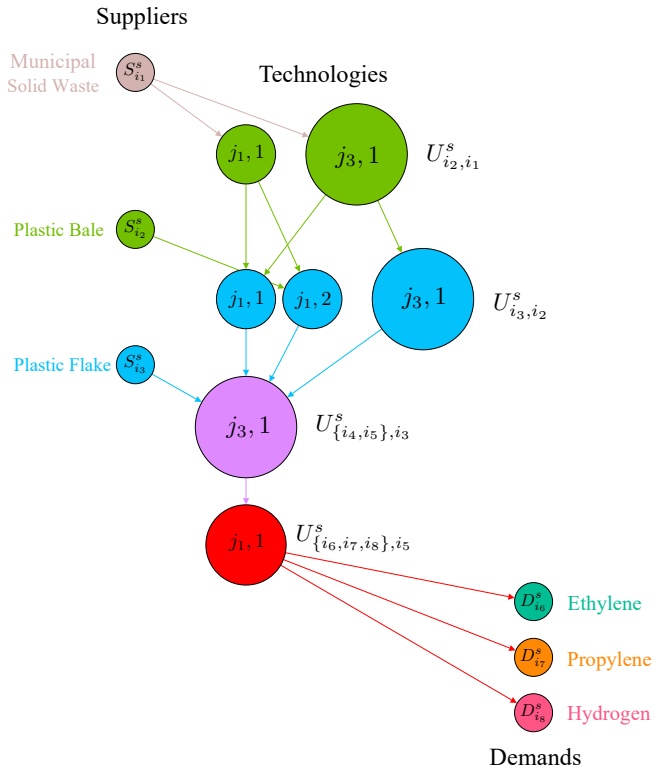


Figure 5.12: Cost-minimizing optimal design of plastic upcycling system (no spatial information).

measure and cost in Table 5.6.

Table 5.5: Module details for optimal system designs with different degrees of modularity.

Modularity Measure, M_4	Annualized Cost, C ($)	Module 1	Module 2	Module 3	Module 4
0.3	6.56×10^8	$U^s_{i_2,i_1,j_1,1}$ $U^s_{i_3,i_2,j_3,1}$	$U^s_{i_2,i_1,j_3,1}$ $U^s_{i_3,i_2,j_1,1}$	$U^s_{i_3,i_2,j_1,2}$ $U^s_{\{i_4,i_5\},i_3,j_3,1}$	$U^s_{\{i_6,i_7,i_8\},i_5,j_1,1}$
0.6	6.95×10^8	$U^s_{i_2,i_1,j_1,1}$ $U^s_{i_2,i_1,j_1,2}$ $U^s_{i_3,i_2,j_1,1}$ $U^s_{i_3,i_2,j_1,2}$	$U^s_{i_2,i_1,j_1,3}$ $U^s_{i_3,i_2,j_1,3}$ $U^s_{\{i_4,i_5\},i_3,j_2,2}$	$U^s_{i_2,i_1,j_2,1}$ $U^s_{\{i_4,i_5\},i_3,j_2,1}$	$U^s_{\{i_4,i_5\},i_3,j_2,1}$ $U^s_{\{i_6,i_7,i_8\},i_5,j_2,1}$

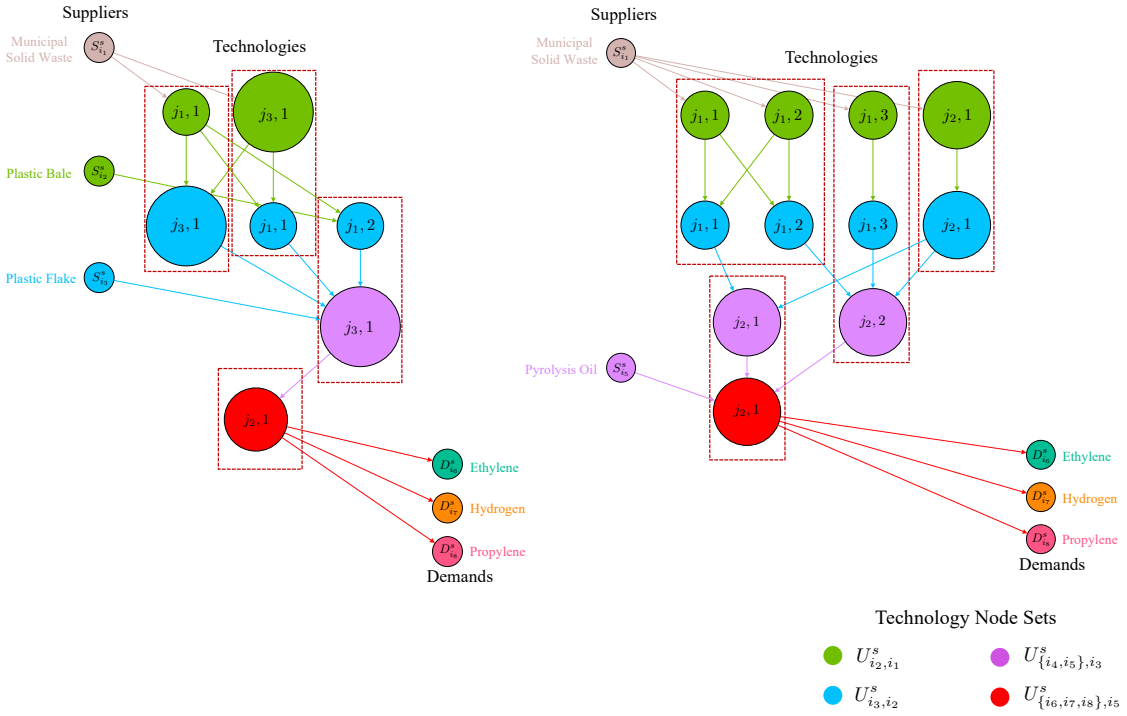


Figure 5.13: Cost-minimizing optimal modular designs with $M_4 \geq 0.3$ (left) and $M_4 \geq 0.6$ (right)

We can see that, as we increase the required degree of modularity, the cost increases and the number of techs installed also increases. This demonstrates that, to achieve a higher modularity measure, smaller techs need to be installed and there exists a trade-off between cost and the degree of modularity of the system. In addition, we observe that, we increase the level of modularity measure from 0 to 0.6, the cost increase is relatively small ($\sim$2% for 0.1 increase in modularity). However, as we further increase the modularity (from 0.6 and above), the cost increases sharply in order to achieve the same amount of modularity increase ($\sim$10% for 0.1 increase in modularity). This means that there exists a threshold for the process design when increasing modularity measure becomes too expensive and makes no economic sense; as such, decision-makers should choose any optimal design within that threshold to achieve a balanced trade-off between cost and modularity. As we increase the modularity measure, the solution time generally becomes

longer since the smaller technologies become relevant here, increasing the possible number of designs and therefore harder for the solver to find an optimal solution. When we are at a high modularity measure (0.8 in this case), the design requirements become too strict, and therefore reduce the number of possible designs available. This explains the reduction on the solution time for the case when the modularity measure is 0.8.

Table 5.6: Trade-offs between system cost and modularity for optimal designs (no spatial information).

Modularity Measure	Cost ($)	# of Technology	Solving Time (s)
0	6.56×10^8	7	0.27
0.3	6.56×10^8	7	1.42
0.4	6.71×10^8	9	22.33
0.5	6.83×10^8	10	16.99
0.6	6.95×10^8	11	22.52
0.7	7.65×10^8	11	48.51
0.8	8.38×10^8	12	17.71

5.4.3 *Optimal System Design with Spatial Information*

When we extend the problem to include spatial information, all the case settings defined in Table 5.2 and Table 5.3 remain valid. We define the additional spatial information such as potential locations to install techs and the transportation cost across different locations in Table 5.7.

Here, a couple of potential locations (B and D) are available to install technologies. Locations for suppliers and demands corresponds to the sequences of product i in set $\mathcal{R}$ and $\mathcal{P}$ as shown in Table 5.3. For example, supply of product i_1 is at location B and demand of product i_6 is at location C. Installation costs for location B and D are represented as a scale times the cost defined in Table 5.3. For instance, installation costs for tech $t \in T_{i_2}$ (technologies T_{i_2,i_1,j_1}, T_{i_2,i_1,j_2}, and T_{i_2,i_1,j_3}) at locations B equals to the scale

Table 5.7: Spatial information for plastic upcycling system.

Parameters	Values
Potential Locations to Install Technologies, $\mathcal{G}_t$	$[B, D]$
Locations for Suppliers of Materials, $\mathcal{G}_s^i, i \in \mathcal{R}$	$[B, C, B, D]$
Locations for Demand of Materials, $\mathcal{G}_d^i, i \in \mathcal{P}$	$[C, D, B, C, B, B]$
Set of All Locations, $\mathcal{G}$	$[B, C, D]$
Scale of Total Cost at $[B, D]$	T_{i_2}: [0.9, 1.1] T_{i_3}: [1, 1.1] $T_{\{i_4,i_5\}}$: [1.1, 1] $T_{\{i_6,i_7,i_8\}}$: [0.9, 1.1]
Transportation Cost, $\alpha^f_{i,B,g'}, i \in \mathcal{I}, g' \in \mathcal{G}$ (\$/unit)	[0.01, 0.1, 0.14]
Transportation Cost, $\alpha^f_{i,C,g'}, i \in \mathcal{I}, g' \in \mathcal{G}$ (\$/unit)	[0.14, 0.01, 0.12]
Transportation Cost, $\alpha^f_{i,D,g'}, i \in \mathcal{I}, g' \in \mathcal{G}$ (\$/unit)	[0.14, 0.11, 0.01]

0.9 times their original installation costs $[0.19, 0.45, 0.89]$ as defined in Table 5.3, and the costs at location D equals to the scale 1.1 times their original costs. Finally, we assume that transport costs for different products are the same but they are different across different locations. For example, transportation cost for any product from location B to D is 0.14 and from location C to D is 0.12. Note that transportation costs at the same location (from location B to B) are much smaller than those across different locations.

The spatial superstructure of the system is shown in Figure 5.14. Nodes on the top represent potential technologies at location B while nodes on the bottom represent their installation at location D. This spatial superstructure is much denser than the superstructure in Figure 5.11; therefore, we have a much larger optimization problem.

We first use formulation (5.3.24) to solve for the cost-minimizing spatial design and the result is shown in Figure 5.15 . This problem has 3253 continuous variables, 50 binary variables, and 3416 constraints, and required 0.16 seconds to solve. The result shows that the optimal design contains 7 techs, with 6 of them being placed at location B while the only pyrolysis plant is placed at location D. This configuration makes sense, as the

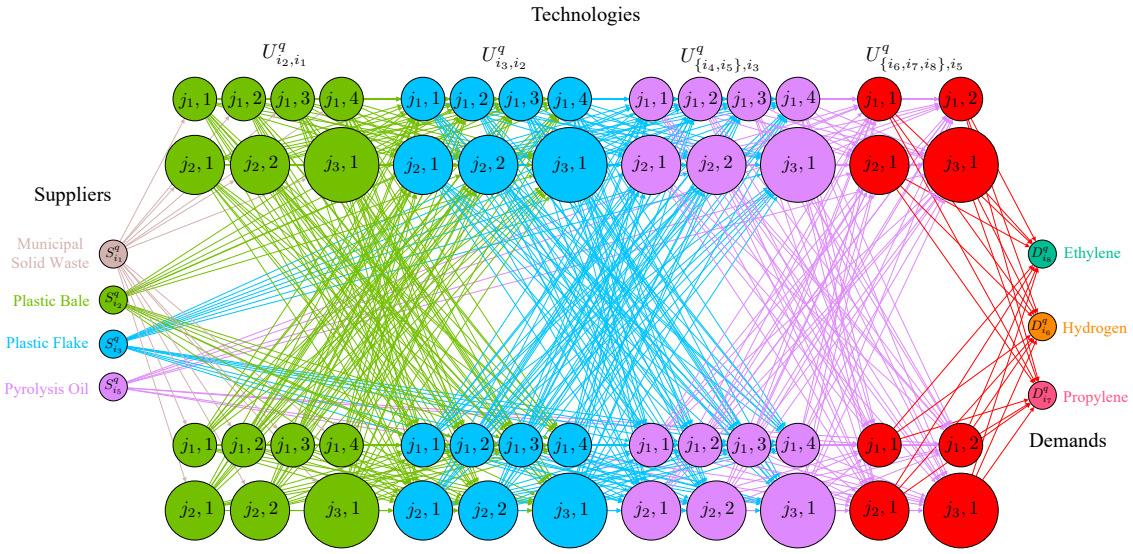


Figure 5.14: Spatial superstructure for plastic waste upcycling system. Technologies on the top are for location B and technologies on the bottom are for location D.

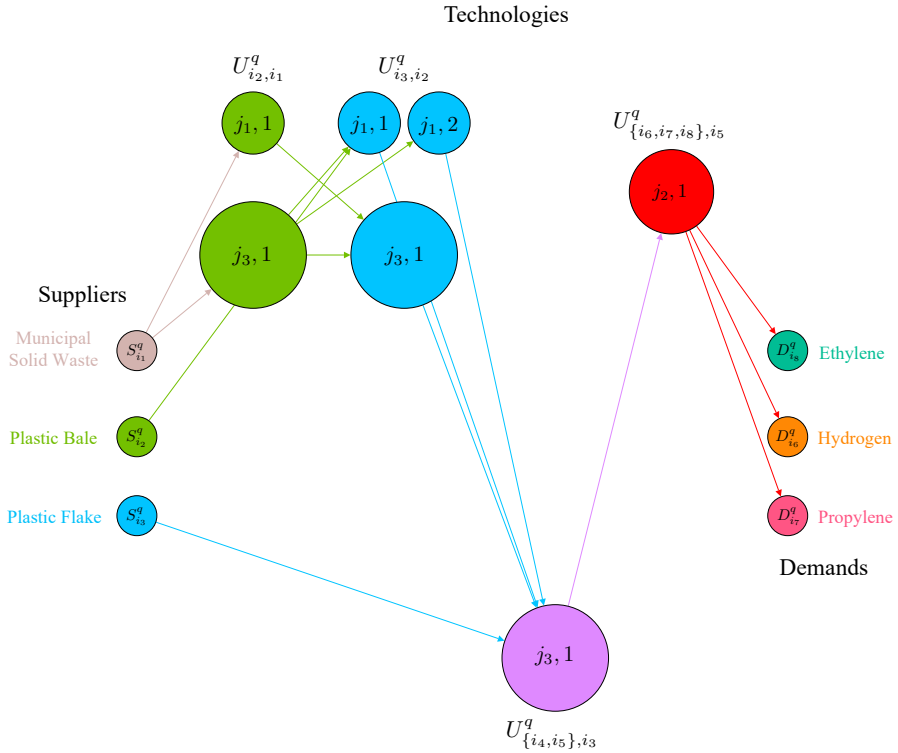


Figure 5.15: Cost-minimizing optimal spatial system design.

optimal design tries to put technologies at locations with less installation cost, which is the major cost of the process. This cost-minimizing design achieves an annualized cost of $\$ 6.30 \times 10^8$ and similarly, it selects the technology with largest capacity for processes MRF, RF, and PP as they are the most cost-efficient units to satisfy the required amount of products. We then use optimization formulation (5.3.28) to solve for the cost-minimizing design with spatial information that also achieves a certain level of modularity.

The optimal designs that correspond to two different levels of modularity are shown in Figure 5.16.

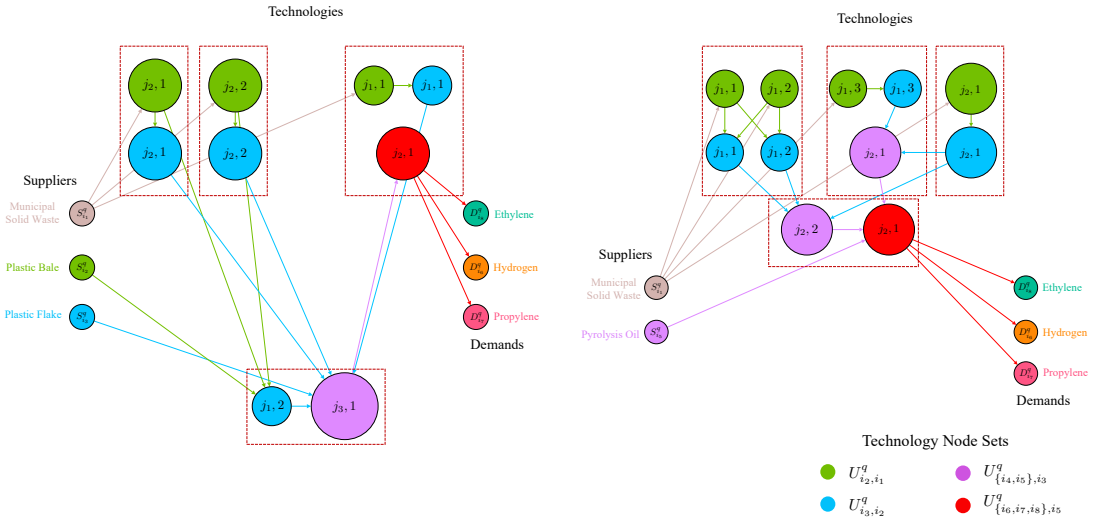


Figure 5.16: Cost-minimizing optimal modular system design (with spatial information) for modularity $M_4 \geq 0.4$ (left) and $M_4 \geq 0.6$ (right)

Module divisions in both cases are grouped by red dashed rectangles and nodes within four modules for each case are summarized in Table 5.8. We can see that, when requirement for modularity is low, the large pyrolysis plant in purple takes the advantage of low installation costs at location D and is grouped with a small reprocessing facility in blue. The module (facility) that these units form is placed at location D, while all other units and modules are placed at a facility in location B. On the right we can see that, as we require a higher degree of modularity, all units and modules are placed at location B. Here, a low installation cost is outweighed by the high transportation cost due to in-

Table 5.8: Details for a couple of modular designs.

Modularity Measure, M_4	Annualized Cost, C ($)	Module 1	Module 2	Module 3	Module 4
0.4	6.46×10^8	$U^q_{i_2,i_1,j_2,1}$ $U^q_{i_3,i_2,j_2,1}$	$U^q_{i_2,i_1,j_2,2}$ $U^q_{i_3,i_2,j_2,2}$	$U^q_{i_2,i_1,j_1,1}$ $U^q_{i_3,i_2,j_1,1}$ $U^q_{\{i_6,i_7,i_8\},i_5,j_2,1}$	$U^q_{i_3,i_2,j_1,2}$ $U^q_{\{i_4,i_5\},i_3,j_3,1}$
0.6	6.74×10^8	$U^q_{i_2,i_1,j_1,1}$ $U^q_{i_2,i_1,j_1,2}$ $U^q_{i_3,i_2,j_1,1}$ $U^q_{i_3,i_2,j_1,2}$	$U^q_{i_2,i_1,j_1,3}$ $U^q_{i_3,i_2,j_1,3}$ $U^q_{\{i_4,i_5\},i_3,j_2,1}$	$U^q_{i_2,i_1,j_2,1}$ $U^q_{i_3,i_2,j_2,1}$	$U^q_{\{i_4,i_5\},i_3,j_2,2}$ $U^q_{\{i_6,i_7,i_8\},i_5,j_2,1}$

creased connectivity across locations and therefore no units are placed at location D. We can also see that smaller techs are utilized in the optimal design with higher modularity. We also summarize the results for different combinations of modularity measure and cost in Table 5.9. A similar trend can be observed in Table 5.6; specifically, a threshold exists around a modularity measure of 0.6 where further increasing the measure causes the cost of design to rise sharply, and the solving time increases dramatically as we increase the requirement for modularity measure initially, and eventually drops when the requirement is too high.

5.5 Conclusions and Future Work

In this work, we propose an optimization framework to facilitate the design of modular manufacturing systems. Central to our approach is the concept of a spatial superstructure, which is a graph that captures all possible system configurations and interdependencies between components. The spatial superstructure is a generalization of the notion of a superstructure and of a p-graph used in process design, in that it encodes spatial (geographical) context of the system components. We show that this generalization enables

Table 5.9: Trade-off between cost and modularity for optimal system design (with spacial information).

Modularity Measure	Cost ($)	# of Technologies	Solving Time (s)
0	6.30×10^8	7	0.92
0.3	6.32×10^8	7	4.89
0.4	6.46×10^8	9	287.51
0.5	6.61×10^8	10	708.48
0.6	6.74×10^8	11	719.71
0.7	7.44×10^8	11	3074.36
0.8	8.17×10^8	12	1312.20

the simultaneous design of processes, facilities, and of supply chains. Our framework aims to select the system topology from the spatial superstructure that minimizes design cost and that maximizes design modularity. We show that this design problem can be cast as a mixed-integer, multi-objective optimization formulation. We demonstrate these capabilities using a case study arising in the design of a plastic waste upcycling supply chain. As part of future work, we are interested in exploring the use of strategies to address computational tractability issues and to capture higher fidelity in the design (e.g., detailed physical models).

6

CONCLUSIONS, CONTRIBUTIONS, AND FUTURE DIRECTIONS

6.1 Contributions

The main contributions of this work are to answer the three questions mentioned in Chapter 1:

1. What - What is modularity? How do we define a quantifiable measure for modularity in manufacturing context?
2. Why - Why should we consider modular design? Can we use mathematical tools to demonstrate the benefits of modular designs comparing to traditional ones?
3. How - How do we design a modular process or supply chain? Can we provide a general framework to aid the synthesis of modular processes and supply chains?

Specifically, they include the development of a quantifiable modularity measure that fits in the manufacturing context, the verification from a mathematical and optimization point of view on the spatial and temporal flexibility brought by modular technologies, and the development of a design scheme that facilitates modular process and supply chain design using computational and analytical tools such as exploratory data analysis, Pareto trade-off analysis, and conflict resolution methodologies. Each newly proposed concept

is accompanied with illustrative visualizations and realistic case studies, and any mathematical demonstrations and optimization formulations are under careful considerations and rigorous derivations.

Quantifiable Modularity Measure in Manufacturing Context

In Chapter 2, we propose a measure to quantify the modularity of industrial production (manufacturing) systems with optimization formulations to compute it. From a manufacturing perspective, we argue that a system is deemed modular if: i) the equipment units that comprise it form clusters (modules) of dense connectivity (i.e., difficult module assembly tasks are performed off-site), ii) connectivity between modules is sparse (i.e., easy assembly tasks are performed on-site), iii) the number of modules is small, and iv) the module dimensions facilitate transportation. We show that the measure proposed satisfies these requirements and that it can be computed by solving a convex mixed-integer quadratic program. We provide a discussion on advantages and disadvantages of alternative modularity measures used in different scientific and engineering communities. Our results seek to highlight conceptual and computational challenges that arise from the need to define and quantify modularity in a manufacturing context. This chapter addresses the "What" question by clarifying the definition of modularity and identifying some unique characteristics of modularity in manufacturing.

Benefits of Modular Design - Spatial Flexibility

In Chapter 3, we talk about the spatial flexibility via technology decentralization in electricity markets. Specifically. We study economic incentives provided by space-time dynamics of day-ahead and real-time electricity markets. We seek to analyze to what extent such dynamics promote decentralization of technologies for generation, consumption, and storage (which is essential to obtain a more flexible power grid). Incentives for decentralization are also of relevance given recent interest in the deployment of small-scale modular technologies (e.g., modular ammonia and biogas production systems). Our

analysis is based on an asset placement problem that seeks to find optimal locations for generators and loads in the network that minimize profit risk. We show that an unconstrained version of this problem can be cast as an eigenvalue problem. Under this representation, optimal network allocations are eigenvectors of the space-time price covariance matrix while the eigenvalues are the associated profit variances. We also construct a more sophisticated placement formulation that captures different risk metrics and constraints on types of technologies to systematically analyze trade-offs in expected profit and risk. Our analysis reveals that space-time market dynamics provide significant incentives for decentralization and strategic asset placement but that full mitigation of risk is only possible through simultaneous investment in generation and loads (which can be achieved using batteries or microgrids).

Benefits of Modular Design - Temporal Flexibility

In Chapter 4, we talk about the temporal flexibility by modular technologies under the setting of a capacity expansion problem. Specifically, we study logistical investment flexibility provided by modular processing technologies for mitigating risk. We propose a multi-stage stochastic programming formulation that determines optimal capacity expansion plans that mitigate demand uncertainty. The formulation accounts for multi-product dependencies between small/large units and for trade-offs between expected profit and risk. The formulation uses a cumulative risk measure to avoid time-consistency issues of traditional, per-stage risk-minimization formulations and we argue that this approach is more compatible with typical investment metrics such as the net present value. Case studies of different complexity are presented to illustrate the developments. Our studies reveal that the Pareto frontier of a flexible setting (allowing for deployment of small units) dominates the Pareto frontier of an inflexible setting (allowing only for deployment of large units). Notably, this dominance is prevalent despite benefits arising from economies of scale of large processing units. Both Chapter 3 and Chapter 4 address the "Why" question.

A Spatial Superstructure Approach to the Optimal Design of Modular Systems

Chapter 5 addresses the "How" question by proposing an optimization framework to facilitate the design of modular manufacturing systems. Central to our approach is the concept of a spatial superstructure, which is a graph that captures all possible system configurations and interdependencies between components. The spatial superstructure is a generalization of the notion of a superstructure and of a p-graph used in process design, in that it encodes spatial (geographical) context of the system components. We show that this generalization enables the simultaneous design of processes, facilities, and of supply chains. Our framework aims to select the system topology from the spatial superstructure that minimizes design cost and that maximizes design modularity. We show that this design problem can be cast as a mixed-integer, multi-objective optimization formulation. We demonstrate these capabilities using a case study arising in the design of a plastic waste upcycling supply chain.

6.2 Future Research Directions

The future research directions are mainly motivated by three aspects of the modularity related problem that our work does not cover, but can be good additions to further understand and extend the modularity concept. First, the complexity of modularity related optimization formulations can grow exponentially with increased resolution, and proper solution approximation methodology should be developed and implemented. Second, the economical assumptions we make about modular technologies need further consideration, and universally acceptable simulation framework for modeling modular technologies need to be established. And finally, the concept of modularity is closely related to that of process intensification, and therefore can be combined with it to further facilitate the design of modular and intensified processes.

Complexity in Modularity Related Formulations

The work in this dissertation includes various chemical process and supply chain case studies under certain assumptions, and they are solved by state-of-art commercial solvers, such as Gurobi. We still observe that some of those models cannot be solved to optimality within reasonable CPU time. This is because the model size explodes exponentially as the resolution (for instance, time resolution can only be 20 minutes in real time electricity markets, and number of locations and potential technologies can not be too large) increases. Furthermore, the mixed-integer formulations of the related optimization models involve the decision of binary variables, which introduces additional difficulty in problem-solving.

While chemical processes and supply chains with limited choice of technologies and locations can be tackled properly, an optimal solution may not be obtainable if we expand the decision set of the problem or increase the model resolution. Therefore, it is necessary to develop methodologies to provide good solutions for large-scale, high-resolution decision-making models. In our models, we realize that binary variables are usually sparse, with limited number of non-zero entries. This characteristics of the model fits into some well-developed decomposition or relaxation methods that can largely help with the problem size. For instance, column generation (Desaulniers et al., 2006; Lübbecke, 2010) and branch-and-price (Barnhart et al., 1998; Savelsbergh, 1997) are popular tools for large-scale integer problems. There are studies on using Dantzig-Wolfe decomposition which is based on the delayed column generation methods for solving multistage stochastic planning problems (Singh et al., 2009), and using Benders decomposition and Lagrangian relaxation as decomposition methods for network design has also been explored (Gendron, 2011; Panconesi and Srinivasan, 1996). As we consider chemical processes and supply chains using a network representation, these methods are well-suited in our case. In pure graph theory, similar problems such as community detection and correlation clustering, have long been studied, and algorithms such as polynomial-time approximation scheme (PTAS) Arora (1998) could also be investigated and applied in our case. In addition, new

methodologies are proposed in recent years that aim to solve similar problems. For instance, semi-definite relaxation methods have been explored to solve large-scale binary quadratic optimization problems that are suitable in solving the modularity measure in our case (Wang et al., 2016), and using unsupervised neural network to solve large-scale multiobjective optimization problem with sparse solutions has also been proposed (Tian et al., 2020). We also note that since there is a trade-off between the model resolution and complexity, methods in selecting optimal resolution for decision-making should be developed and validated.

Assumptions in Modeling Modular Technologies and Processes

For the case studies that performed in this work, modular technologies are usually modeled as ones with smaller capacity with a trade-off of more expensive per unit cost based on the economy of scale. To make things simple, the capital cost of a modular technology is assumed to be based on the "six-tenth" rule (or similarly "two-thirds" rule). However, the differences between modular technology and conventional facility are much more profound, and sometimes cannot be evaluated economically. For instance, modular technologies are usually shop fabricated that are not delayed by weather conditions. This makes the project length much more manageable, and can potentially save on construction cost. Also, as multiple copies of the same modular technologies are often needed to expand the capacity, the workers become sophisticated during the process which can also reduce the project length and save on the building cost. As modular technologies are usually built with movability, the cost of transporting, assembling, and disassembling modules should also be accurately estimated and included. From another perspective, a smaller construction crew and the indoor building environment can potentially reduce the carbon footprint of the project. As more and more companies are now making bold declarations about their carbon neutral plans, this advantage of modular technologies, even though not measurable directly by money, are even more important nowadays comparing to some percentage cost savings. As going modular is still a relatively new concept

in the industry, there are currently no broadly accepted rules of thumbs like the "six-tenth" rule on estimating the cost factors of modular technologies. Researchers have looked into estimating the fixed capital investment for modular production plants, and a net present value analysis has been done on modular chemical plants (Lier and Grünewald, 2011; Sievers et al., 2017). However, there is still a lack of contributions in this research field, and as one of the fundamental assumptions that my work depends on, establishing a rule for modeling modular technologies can be an important and interesting future direction.

Connections with Process Intensification

Process intensification is another popular concept that tends to make dramatic reductions in the size of a chemical plant (Stankiewicz and Moulijn, 2002; Stankiewicz et al., 2000; Van Gerven and Stankiewicz, 2009). The concepts of modular technologies and process intensification are usually mentioned together in literature (Bielenberg and Palou-Rivera, 2019; Kim et al., 2017b), but unlike modular technologies which can achieve size reduction simply by scaling down of a process, process intensification requires combinations of multiple unit operations to achieve size reduction and usually involves development of new technologies and operating equipment such as membrane reactors (Becht et al., 2009; Drioli et al., 2011). Therefore, an intensified process is generally harder to design and can vary for different chemical processes. However, performing modular analysis for a conventional process can help identify unit operations that are closely tied together and therefore more likely to be intensified. In general, modular technology and intensified process are closely related, and investigating how to combine these two concepts such as using modular analysis to help identify potentially "intensifiable" operations is an interesting future direction.

ACKNOWLEDGEMENT

We acknowledge financial support from the National Science Foundation (CAREER award grant CBET-1748516), and partial support from the U.S. Department of Agriculture (grant 2017-67003-26055).

BIBLIOGRAPHY

Federal size regulations for commercial motor vehicles. https://ops.fhwa.dot.gov/freight/publications/size_regs_final_rpt/index.htm, 2017. last visit on 2022-01.

Agarwal, G. and Kempe, D. Modularity-maximizing graph communities via mathematical programming. *The European Physical Journal B*, 66:409–418, Dec 2008.

Allman, A. and Zhang, Q. Dynamic location of modular manufacturing facilities with relocation of individual modules. *European Journal of Operational Research*, 2020.

Allman, A., Tang, W., and Daoutidis, P. Decode: a community-based algorithm for generating high-quality decompositions of optimization problems. *Optimization and Engineering*, 20(4):1067–1084, 2019.

AM and EC. Interface specificaiton for OASIS (spring 2014/ferc764 release), 2014. [Accessed July 2, 2018].

Arora, S. Polynomial time approximation schemes for euclidean traveling salesman and other geometric problems. *Journal of the ACM*, 45(5):753–782, 1998.

Avgerinou, M., Bertoldi, P., and Castellazzi, L. Trends in data centre energy consumption under the european code of conduct for data centre energy efficiency. *Energies*, 10:1470, 2017.

Baldea, M., Edgar, T. F., Stanley, B. L., and Kiss, A. A. Modular manufacturing processes: Status, challenges, and opportunities. *AIChE Journal*, 63(10):4262–4272, 2017.

Baldwin, C. Y., Clark, K. B., Clark, K. B., et al. *Design rules: The power of modularity*, volume 1. MIT press, 2000.

Barnhart, C., Johnson, E. L., Nemhauser, G. L., Savelsbergh, M. W., and Vance, P. H. Branch-and-price: Column generation for solving huge integer programs. *Operations research*, 46(3):316–329, 1998.

Becht, S., Franke, R., Geißelmann, A., and Hahn, H. An industrial view of process intensification. *Chemical Engineering and Processing: Process Intensification*, 48(1):329–332, 2009.

Becker, T., Bruns, B., Lier, S., and Werners, B. Decentralized modular production to increase supply chain efficiency in chemical markets. *Journal of Business Economics*, 91 (6):867–895, 2021.

Beddoes, J., Bracmort, K., Burns, R., and Lazarus, W. An analysis of energy production costs from anaerobic digestion systems on u.s. livestock production facilities. Technical Report Technical Note No.1, United States Department of Agriculture, 10 2007.

Berthouex, P. M. Evaluating economy of scale. *Water Pollution Control Federation*, 44(11): 2111–2119, 1972.

Berthélemy, M. and Rangel, L. E. Nuclear reactors' construction costs: The role of lead-time, standardization and technological progress. *Energy Policy*, 82:118 – 130, 2015.

Bhattacharyya, D., Shaeiwitz, J. A., Whiting, W. B., Bailie, R. C., and Turton, R. *Analysis, Synthesis, and Design of Chemical Processes*. Prentice Hall, 2012.

Bhosekar, A., Badejo, O., and Ierapetritou, M. Modular supply chain optimization considering demand uncertainty to manage risk. *AIChE Journal*, 67(11):e17367, 2021.

Bielenberg, J. and Palou-Rivera, I. The rapid manufacturing institute–reenergizing us efforts in process intensification and modular chemical processing. *Chemical Engineering and Processing-Process Intensification*, 138:49–54, 2019.

Bieringer, T., Buchholz, S., and Kockmann, N. Future production concepts in the chemical industry: modular–small-scale–continuous. *Chemical Engineering & Technology*, 36(6): 900–910, 2013.

Boda, K. and Filar, J. Time consistent dynamic risk measures. *Math Meth Oper Res*, 63: 169–186, 2006.

Bramsiepe, C., Sievers, S., Seifert, T., Stefanidis, G., Vlachos, D. G., Schnitzer, H., Muster, B., Brunner, C., Sanders, J., Bruins, M., et al. Low-cost small scale processing technologies for production applications in various environmentsÑmass produced factories. *Chemical Engineering and Processing: Process Intensification*, 51:32–52, 2012.

Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hoefer, M., Nikoloski, Z., and Wagner, D. On modularity clustering. *IEEE transactions on knowledge and data engineering*, 20(2): 172–188, 2007.

Buchholz, S. Future manufacturing approaches in the chemical and pharmaceutical industry. *Chemical Engineering and Progressing: Process Intensification*, 49(10):993–995, 2010.

Cabe, W. M., Smith, J., Harriott, P., et al. *Unit Operation of Chemical Engineering*. McGraw-Hill, 2018.

Cardin, M.-A. and Hu, J. Analyzing the Tradeoffs Between Economies of Scale, Time-Value of Money, and Flexibility in Design Under Uncertainty: Study of Centralized Versus Decentralized Waste-to-Energy Systems. *Journal of Mechanical Design*, 138(1), 11 2015.

Chakraborty, S., Kramer, B., and Kroposki, B. A review of power electronics interfaces for distributed energy systems towards achieving low-cost modular design. *Renewable and Sustainable Energy Reviews*, 13(9):2323 – 2335, 2009.

Charpentier, J.-C. The triplet "molecular processes–product–process" engineering: the future of chemical engineering? *Chemical Engineering Science*, 57(22-23):4667–4690, 2002.

Chen, Q. and Grossmann, I. E. Economies of numbers for a modular stranded gas processing network: Modeling and optimization. In *Computer Aided Chemical Engineering*, volume 47, pages 257–262. Elsevier, 2019.

Chiang, N.-Y. and Zavala, V. M. Large-scale optimal control of interconnected natural gas and electrical transmission systems. *Applied Energy*, 168:226 – 235, 2016.

Christodoulos A Floudas, A. R. C. and Grossmann, I. E. Automatic synthesis of optimum heat exchanger network configurations. *AIChE Journal*, 32(2):276–290, 1986.

Clauset, A., Newman, M. E., and Moore, C. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004.

Coltheart, M. Modularity and cognition. *Trends in cognitive sciences*, 3(3):115–120, 1999.

Conejo, A., Fernandez-Gonzalez, J., and Alguacil, N. Energy procurement for large consumers in electricity markets. *IEE Proceedings-Generation, Transmission and Distribution*, 152(3):357–364, 2005.

Constantino, P. H., Tang, W., and Daoutidis, P. Topology effects on sparse control of complex networks with laplacian dynamics. *Scientific reports*, 9(1):9034, 2019.

Daoutidis, P., Tang, W., and Jogwar, S. S. Decomposing complex plants for distributed control: Perspectives from network theory. *Computers & Chemical Engineering*, 114:43–51, 2018.

Daoutidis, P., Tang, W., and Allman, A. Decomposition of control and optimization problems by network structure: Concepts, methods, and inspirations from biology. *AIChE Journal*, 65(10):e16708, 2019.

DapkLus, W. D. and Bowe, T. R. Planning for new electric generation technologies a stochastic dynamic programming approach. *IEEE Transactions on Power Apparatus and Systems*, PAS-103(6):1447–1453, 1984.

Davis, N. Natural gas flaring in north dakota has declined sharply since 2014. https://www.eia.gov/todayinenergy/detail.php?id=26632, 2016. last visit on 2022-01.

Desaulniers, G., Desrosiers, J., and Solomon, M. M. *Column generation*, volume 5. Springer Science & Business Media, 2006.

Donath, W. E. and Hoffman, A. J. Lower bounds for the partitioning of graphs. *IBM J. Res. Dev.*, 17(5):420–425, September 1973.

Dong, L., Liu, H., and Riffat, S. Development of small-scale and micro-scale biomass-fuelled chp systems – a literature review. *Applied Thermal Engineering*, 29(11):2119 – 2126, 2009.

Dong, Y. and Qin, S. J. A novel dynamic pca algorithm for dynamic data modeling and process monitoring. *Journal of Process Control*, 67:1 – 11, 2018.

Doran, D. and Giannakis, M. An examination of a modular supply chain: a construction sector perspective. *Supply Chain Management: An International Journal*, 2011.

Dowling, A. W., Kumar, R., and Zavala, V. M. A multi-scale optimization framework for electricity market participation. *Applied Energy*, 190:147–164, 2017.

Drioli, E., Stankiewicz, A. I., and Macedonio, F. Membrane engineering in process intensification—an overview. *Journal of Membrane Science*, 380(1-2):1–8, 2011.

Egenhofer, C., Schrefler, L., Rizos, V., Infelise, F., Luchetta, G., Simonelli, F., Stoefs, W., Timini, J., and Colantoni, L. Composition and drivers of energy prices and costs in energy-intensive industries: The case of the chemical industry-ammonia, 2017. last visit on 2022-01.

El-Halwagi, M. M. and Manousiouthakis, V. Synthesis of mass exchange networks. *AIChE Journal*, 35(8):1233–1244, 1989.

Emam, E. A. Gas flaring in industry: An overview. *Petroleum & coal*, 57(5), 2015.

Epic Systems. Keys to modular chemical plants. last visit on 2022-01.

F Friedler, Y. H., K Tarjan and Fan, L. Graph-theoretic approach to process synthesis: axioms and theorems. *Chemical Engineering Science*, 47(8):1973–1988, 1992.

F.H. Murphy, S. S. and Soyster, A. Electric utility capacity expansion planning with uncertain load forecasts. *IIE Trans.*, 14:52–59, 1982.

Fisher, M. L. The lagrangian relaxation method for solving integer programming problems. *Manage. Sci.*, 50(12 Supplement):1861–1871, dec 2004.

Fortunato, S. Community detection in graphs. *Physics Reports*, 486(3):75 – 174, 2010.

Frivaldsky, M., Spanik, P., Morgos, J., and Pridala, M. Control strategy proposal for modular architecture of power supply utilizing lcct converter. *Energies*, 11(12):3327, 2018.

Gendron, B. Decomposition methods for network design. *Procedia-Social and Behavioral Sciences*, 20:31–37, 2011.

Geng, N., Jiang, Z., and Chen, F. Stochastic programming based capacity planning for semiconductor wafer fab with uncertain demand and capacity. *European Journal of Operational Research*, 198(3):899 – 908, 2009.

Georgia Power. Hammond plant. [Accessed July 23, 2018].

Guo, C., Lu, G., Li, D., Wu, H., Zhang, X., and Shi, Y. Bcube: A high performance, server-centric network architecture for modular data centers. *ACM SIGCOMM Computer Communication Review*, 39(4):63–74, 2009.

Gwehenberger, G. and Narodoslawsky, M. Sustainable processes—the challenge of the 21st century for chemical engineering. *Process safety and environmental protection*, 86(5): 321–327, 2008.

Hady, Ł., Dyląg, M., and Wozny, G. Investment cost estimation and calculation of chemical plants with classical and modular approaches. *Chemical and Process Engineering*, 30: 319–340, 2009.

Hagspiel, V., Huisman, K. J., and Kort, P. M. Volume flexibility and capacity investment under demand uncertainty. *International Journal of Production Economics*, 178:95–108, 2016.

Haney, F., Donovan, G., Roth, T., Lowrie, A., Morlidge, G., Lucchini, S., and Halvorsen, S. Modular processing facility.

Held, M. and Karp, R. M. The traveling-salesman problem and minimum spanning trees: Part ii. *Mathematical Programming*, 1(1):6–25, Dec 1971.

Heo, S., Marvin, W. A., and Daoutidis, P. Automated synthesis of control configurations for process networks based on structural coupling. *Chemical Engineering Science*, 136: 76–87, 2015.

Hesler, W. E. Modular design - where it fits. *Chemical Engineering Process*, October:76–80, 1990.

Heuberger, C. F., Rubin, E. S., Staffell, I., Shah, N., and Dowell, N. M. Power capacity expansion planning considering endogenous technology cost learning. *Applied Energy*, 204:831 – 845, 2017.

Hohmann, L., Kössl, K., Kockmann, N., Schembecker, G., and Bramsiepe, C. Modules in process industry - a life cycle definition. *Chemical Engineering and Processing: Process Intensification*, 111:115 – 126, 2017.

Hu, Y., Scarborough, M., Aguirre-Villegas, H., Larson, R. A., Noguera, D. R., and Zavala, V. M. A supply chain framework for the analysis of the recovery of biogas and fatty acids from organic waste. *ACS Sustainable Chemistry & Engineering*, 6(5):6211–6222, 2018.

Isafiade, A. and Fraser, D. Interval based minlp superstructure synthesis of combined heat and mass exchanger networks. *Chemical Engineering Research and Design*, 87(11): 1536–1542, 2009.

ISO New England. Operating the power system. [Accessed July 23, 2018].

Jaikumar, R. Postindustrial manufacturing. *Harvard Business Review*, 64(6):69–76, 1986.

Jogwar, S. S. and Daoutidis, P. Community-based synthesis of distributed control architectures for integrated process networks. *Chemical Engineering Science*, 172:434–443, 2017.

Johnson, E. P. and Oliver, M. E. Renewable energy and wholesale electricity price variable. In *Proceedings of the 1st Quarter 2016 IAEE Energy Forum*, pages 25–26, 2016.

Kerridge, A. Evaluate project cost factors. *Hydrocarbon Process.;(United States)*, 61(7), 1982.

Kim, K., Yang, F., Zavala, V. M., and Chien, A. A. Data centers as dispatchable loads to harness stranded power. *IEEE Transactions on Sustainable Energy*, 8(1):208–218, 2017a.

Kim, Y.-h., Park, L. K., Yiacoumi, S., and Tsouris, C. Modular chemical process intensification: a review. *Annual review of chemical and biomolecular engineering*, 8:359–380, 2017b.

Kockmann, N. Modular equipment for chemical process development and small-scale production in multipurpose plants. *ChemBioEng Reviews*, 3(1):5–15, 2016.

Kockmann, N., Thenée, P., Fleischer-Trebes, C., Laudadio, G., and Noël, T. Safety assessment in development and operation of modular continuous-flow processes. *Reaction Chemistry & Engineering*, 2(3):258–280, 2017.

Krich, K., Augenstein, D., Batmale, J., Benemann, J., Rutledge, B., and Salour, D. Biomethane from dairy waste: A sourcebook for the production and use of renewable natural gas in california. Technical report, Western United Dairymen, 07 2005.

Langlois, R. N. Modularity in technology and organization. *Journal of economic behavior & organization*, 49(1):19–37, 2002.

Ledoit, O. and Wolf, M. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of multivariate analysis*, 88(2):365–411, 2004.

Lier, S. and Grünewald, M. Net present value analysis of modular chemical production plants. *Chemical Engineering & Technology*, 34(5):809–816, 2011.

Lier, S., Wörsdörfer, D., and Grünewald, M. Transformable production concepts: Flexible, mobile, decentralized, modular, fast. *ChemBioEng Reviews*, 3(1):16–25, 2016.

Liu, Y., Sioshansi, R., and Conejo, A. J. Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. *IEEE Transactions on Power Systems*, 33(1):781–791, 2018.

Lübbecke, M. E. Column generation. *Wiley encyclopedia of operations research and management science. Wiley, New York*, pages 1–14, 2010.

Luss, H. Capacity expansion model for two facility types. *Naval Research Logistics Quarterly*, 26(2):291–301, 1979.

Luss, H. Multi-facility capacity expansion model with joint expansion set-up costs. *Naval Research Logistics Quarterly*, 30(1):97–112, 1983.

Luss, H. Heuristic for capacity planning with multiple facility types. *Naval Research Logistics Quarterly*, 33(4):685–702, 1986.

Meunier, D., Lambiotte, R., Fornito, A., Ersche, K. D., and Bullmore, E. T. Hierarchical modularity in human brain functional networks. *Frontiers in neuroinformatics*, 3, 2009.

Meunier, D., Lambiotte, R., and Bullmore, E. T. Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, 4:200, 2010.

Moharir, M., Kang, L., Daoutidis, P., and Almansoori, A. Graph representation and decomposition of ode/hyperbolic pde systems. *Computers & Chemical Engineering*, 106: 532–543, 2017.

Moharir, M., Pourkargar, D. B., Almansoori, A., and Daoutidis, P. Graph representation and distributed control of diffusion-convection-reaction system networks. *Chemical Engineering Science*, 204:128–139, 2019.

Newman, M. E. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.

openstax. Principles of economics. last visit on 2022-01.

Palys, M. J., Allman, A., and Daoutidis, P. Exploring the benefits of modular renewable-powered ammonia production: A supply chain optimization study. *Industrial & Engineering Chemistry Research*, 2018.

Panconesi, A. and Srinivasan, A. On the complexity of distributed network decomposition. *Journal of Algorithms*, 20(2):356–374, 1996.

Patel, P. How inexpensive must energy storage be for utilities to switch to 100 percent renewables? https://spectrum.ieee.org/energywise/energy/renewables/what-energy-storage-would-have-to-cost-for-a-renewable-grid, 2019. Accessed: 2020-07-05.

Peters, M. S., Timmerhaus, K. D., West, R. E., Timmerhaus, K., and West, R. *Plant design and economics for chemical engineers*, volume 4. McGraw-Hill New York, 1968.

Pfaudler Group. Modular reaction, evaporation & distillation system. last visit on 2022-01.

Pionner Energy. Mobile flare gas capture solutions & modular gas processing plants. last visit on 2022-01.

Pourkargar, D. B., Moharir, M., Almansoori, A., and Daoutidis, P. Distributed estimation and nonlinear model predictive control using community detection. *Industrial & Engineering Chemistry Research*, 58(30):13495–13507, 2019.

Pourkargar, D. B., Almansoori, A., and Daoutidis, P. Comprehensive study of decomposition effects on distributed output tracking of an integrated process over a wide operating range. *Chemical Engineering Research and Design*, 134:553–563, 2018.

R., D. Small geothermal power plants: design, performance and economics. *GHC bulletin*, 20(2), 1999.

Rahman, S. M. T., Salim, M. T., and Syeda, S. R. Facility layout optimization of an ammonia plant based on risk and economic analysis. *Procedia Engineering*, 90:760–765, 2014.

Rajagopalan, S. Flexible versus dedicated technology: A capacity expansion model. *International Journal of Flexible Manufacturing Systems*, 5(2):129–142, 1993.

Ramshaw, C. Process intensification and green chemistry. *Green Chemistry*, 1:G15–G17, 1999.

Rattigan, M. J., Maier, M., and Jensen, D. Graph clustering with network structure indices. In *Proceedings of the 24th International Conference on Machine Learning*, ICML '07, pages 783–790, New York, NY, USA, 2007. ACM.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A.-L. Hierarchical organization of modularity in metabolic networks. *science*, 297(5586):1551–1555, 2002.

Reichardt, J. and Bornholdt, S. Statistical mechanics of community detection. *Physical review E*, 74(1):016110, 2006.

Rogers, G. and Bottaci, L. Modular production systems: a new manufacturing paradigm. *Journal of intelligent manufacturing*, 8(2):147–156, 1997.

Rosvall, M. and Bergstrom, C. T. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, 105(4):1118–1123, 2008.

Rothwell, G. A real options approach to evaluating new nuclear power plants. *The Energy Journal*, 27(1):37–53, 2006.

Roy, S. and Eng, P. Consider modular plant design. *Chemical Engineering Progress*, 113(5): 28–31, 2017.

Sampat, A. M., Ruiz-Mercado, G. J., and Zavala, V. M. Economic and environmental analysis for advancing sustainable management of livestock waste: A wisconsin case study. *ACS Sustainable Chemistry & Engineering*, 6(5):6018–6031, 2018.

Savelsbergh, M. A branch-and-price algorithm for the generalized assignment problem. *Operations research*, 45(6):831–841, 1997.

Seifert, T., Sievers, S., Bramsiepe, C., and Schembecker, G. Small scale, modular and continuous: a new approach in plant design. *Chemical Engineering and Processing: Process Intensification*, 52:140–150, 2012.

Seliger, G. and Zettl, M. Modularization as an enabler for cycle economy. *CIRP Annals-Manufacturing Technology*, 57(1):133–136, 2008.

Shao, Y. and Zavala, V. M. Space-time dynamics of electricity markets incentivize technology decentralization. *Computers & Chemical Engineering*, 127:31–40, 2019.

Shao, Y. and Zavala, V. M. Modularity measures: Concepts, computation, and applications to manufacturing systems. *AIChE Journal*, 66(6):e16965, 2020.

Shao, Y., Hu, Y., and Zavala, V. M. Mitigating investment risk using modular technologies. *Computers & Chemical Engineering*, 153:107424, 2021.

Shiina, T. and Birge, J. R. Multistage stochastic programming model for electric power capacity expansion problem. *Japan Journal of Industrial and Applied Mathematics*, 20(3): 379, 2003.

SHIINA, T., TAKAICHI, T., LI, Y., MORITO, S., and IMAIZUMI, J. Multistage stochastic programming model and solution algorithm for the capacity expansion of railway network. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 12 (3):JAMDSM0077, 2018.

Sievers, S., Seifert, T., Franzen, M., Schembecker, G., and Bramsiepe, C. Fixed capital investment estimation for modular production plants. *Chemical Engineering Science*, 158:395–410, 2017.

Silvestre, J. Economies and diseconomies of scale. *The New Palgrave: A Dictionary of Economics*, 2:80–84, 1987.

Simon, H. A. The architecture of complexity. *Proceedings of the American philosophical society*, 106(6):467–482, 1962.

Singh, K. J., Philpott, A. B., and Wood, R. K. Dantzig-wolfe decomposition for solving multistage stochastic capacity-planning problems. *Operations Research*, 57(5):1271–1286, 2009.

Sioshansi, R., Denholm, P., Jenkin, T., and Weiss, J. Estimating the value of electricity storage in pjm: Arbitrage and some welfare effects. *Energy economics*, 31(2):269 – 277, 2009.

Soltanieh, M., Zohrabian, A., Gholipour, M. J., and Kalnay, E. A review of global gas flaring and venting and impact on the environment: Case study of iran. *International Journal of Greenhouse Gas Control*, 49:488–509, 2016.

Stankiewicz, A. and Moulijn, J. A. Process intensification. *Industrial & Engineering Chemistry Research*, 41(8):1920–1924, 2002.

Stankiewicz, A. I., Moulijn, J. A., et al. Process intensification: transforming chemical engineering. *Chemical engineering progress*, 96(1):22–34, 2000.

Stevens, M. Chemical engineering: The practice of the profession "pace, price, perplexities". In *Chemical engineering: Visions of the World*, pages 91–103. Elsevier, 2003.

Sun, Y. and Schonfeld, P. Stochastic capacity expansion models for airport facilities. *Transportation Research Part B: Methodological*, 80:1 – 18, 2015.

Tang, W., Pourkargar, D. B., and Daoutidis, P. Relative time-averaged gain array (rtaga) for distributed control-oriented network decomposition. *AIChE Journal*, 64(5):1682–1690, 2018.

Tatum, C. Improving constructibility during conceptual planning. *Journal of Construction Engineering and Management*, 113(2):191–207, 1987.

Tian, Y., Lu, C., Zhang, X., Tan, K. C., and Jin, Y. Solving large-scale multiobjective optimization problems with sparse optimal solutions via unsupervised neural networks. *IEEE transactions on cybernetics*, 51(6):3115–3128, 2020.

Trevor Hastie, R. T. and Friedman, J. *The Elements of Statistical Learning*. Springer-Verlag New York, 2009.

Udok, U. and Akpan, E. B. Gas flaring in nigeria: Problems and prospects. *Global Journal of Politics and Law Research*, 5(1):16–28, 2017.

Van Gerven, T. and Stankiewicz, A. Structure, energy, synergy, time - the fundamentals of process intensification. *Industrial & engineering chemistry research*, 48(5):2465–2474, 2009.

Vanhatalo, E., Kulahci, M., and Bergquist, B. On the structure of dynamic principal component analysis used in statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems*, 167:1 – 11, 2017.

Walker, W. H., Lewis, W. K., and McAdams, W. H. *Principles of chemical engineering*. McGraw-Hill, 1923.

Wang, B. and Hobbs, B. F. A flexible ramping product: Can it help real-time dispatch markets approach the stochastic dispatch ideal? *Electric Power Systems Research*, 109: 128–140, 2014.

Wang, P., Shen, C., van den Hengel, A., and Torr, P. H. Large-scale binary quadratic optimization using semidefinite relaxation and applications. *IEEE transactions on pattern analysis and machine intelligence*, 39(3):470–485, 2016.

Wells, D. D. Movement key to pre-fab module use. *Oil and Gas Journal*, 77(42):148–168, 1979.

Wiendahl H.-P., K. J., Nofen D. and F., B. *Planung Modularer Fabriken*. Carl Hanser Verlag GmbH & Co. KG, 2005.

Wu, H., Lu, G., Li, D., Guo, C., and Zhang, Y. Mdcube: A high performance network structure for modular data center interconnection. In *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies*, pages 25–36, 2009.

Xu, G., Tsoka, S., and Papageorgiou, L. Finding community structures in complex networks using mixed integer optimisation. *The European Physical Journal B*, 60:231–239, 2007.

Yeomans, H. and Grossmann, I. E. A systematic modeling framework of superstructure optimization in process synthesis. *Computers and Chemical Engineering*, 23(6):709–731, 1999.

Zavala, V. M., Kim, K., Anitescu, M., and Birge, J. A stochastic electricity market clearing formulation with consistent pricing properties. *Operations Research*, 65(3):557–576, 2017.

Zhao, S., Haskell, W. B., and Cardin, M.-A. Decision rule-based method for flexible multi-facility capacity expansion problem. *IISE Transactions*, 50(7):553–569, 2018.

Zhao, S., Haskell, W. B., and Cardin, M.-A. A flexible multi-facility capacity expansion problem with risk aversion, 2019.