

Integrated Redundancy and Storage Design Optimization for Reliable Air Separation Units Based on Markov Chain—A Game Theoretic Solution

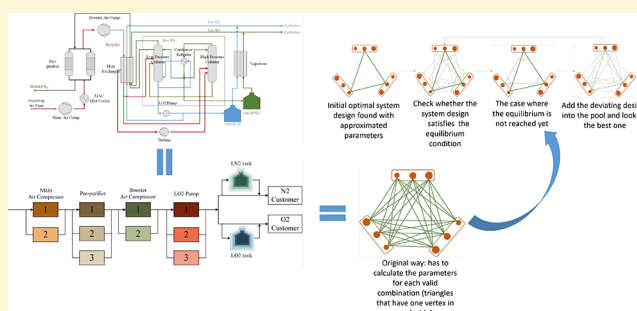
Yixin Ye,[†] Ignacio E. Grossmann,^{*,†} Jose M. Pinto,[‡] and Sivaraman Ramaswamy[‡]

[†]Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15232, United States

[‡]Business and Supply Chain Optimization, Linde.Digital, Linde plc, Tonawanda, New York 14150, United States

Supporting Information

ABSTRACT: The profitability of a chemical plant is directly related to its reliability, which has always been a major concern in the chemical industry. In this paper, we address the problem at the conceptual design phase of an air separation plant to minimize the negative income, which consists of the penalty incurred from pipeline supply interruption and the cost of increasing reliability, including having redundant units and storage tanks. A mixed integer linear programming (MILP) model (denoted as RST) based on the Markov Chain assumption is proposed and applied to the motivating example of an air separation plant. Furthermore, to tackle larger superstructures, we propose a game theoretic algorithm that decomposes and restructures the problem as a team game of the individual processing stages and arrives at a Nash Equilibrium among them. It is also shown that a good initialization point close to the global optimum can be easily obtained, which guarantees the quality of the Nash Equilibrium solution. A number of examples are shown to illustrate the proposed algorithm's ability to solve to global optimality in much shorter time than the direct solution of the original MILP model (RST).



1. INTRODUCTION

The reliability of a chemical plant is very important as it directly impacts the customer service level and its economic performance. A plant with all the state-of-art technologies may not be able to generate the expected profits because of poor decision making regarding its reliability. In industrial practice, significant emphasis is placed on carrying out improved maintenance operations. There has also been academic research on quantifying and optimizing the maintenance efforts after the commissioning of a plant.^{1–5}

However, prior to operations, the decisions regarding system reliability are equally worthy of consideration in the conceptual design phase. Thomaidis et al.^{6,7} integrate flexibility and reliability in process design by deciding the reliability index of each piece of equipment. Gong et al.⁸ incorporate several strategies in both design and operations toward more resilient chemical processes, including a major failure proof strategy for process plants, which is to have redundancies for critical equipment. There is a clear trade-off when higher expected availability of the process requires greater capital investment. Kuo et al.⁹ provide a literature survey on optimal reliability design methods classified in terms of problem formulations and optimization techniques. Aguilar et al.¹⁰ address reliability in utility plant design and operation by considering a few prespecified alternatives for redundancy selection, and for which they formulate an MILP model considering certain failure scenarios. Ye et al.¹¹ propose

a general mixed-integer programming framework for the optimal selection of redundant units.

Another widely used strategy to improve product availability is providing buffer storage for intermediate or final products, which again incurs costs of building the tanks and maintaining the stock that increases with the size of the storage tanks. In our earlier effort¹² to optimize redundancy selection together with maintenance policies, Markov Chain is used to model the stochastic process of failures and repairs. This paper inherits the Markov Chain framework and takes a step back from unifying design and operations decisions to focus on expanding the scope of design decisions with the inclusion of storage tanks. For the ASU processes that motivated this work, it is especially important and common to have storage tanks of liquid products as the last line of defense because the penalty of interrupting pipeline supplies can be very large. In a similar previous work, Terrazas et al.¹³ optimize the reliable design of chemical sites networks with Markov Chain framework, and estimate the stochastic variable of a buffer storage level with the linear upper bound of its variance. As an improvement, this

Special Issue: Christos Georgakis Festschrift

Received: August 20, 2019

Revised: November 5, 2019

Accepted: November 27, 2019

Published: November 27, 2019



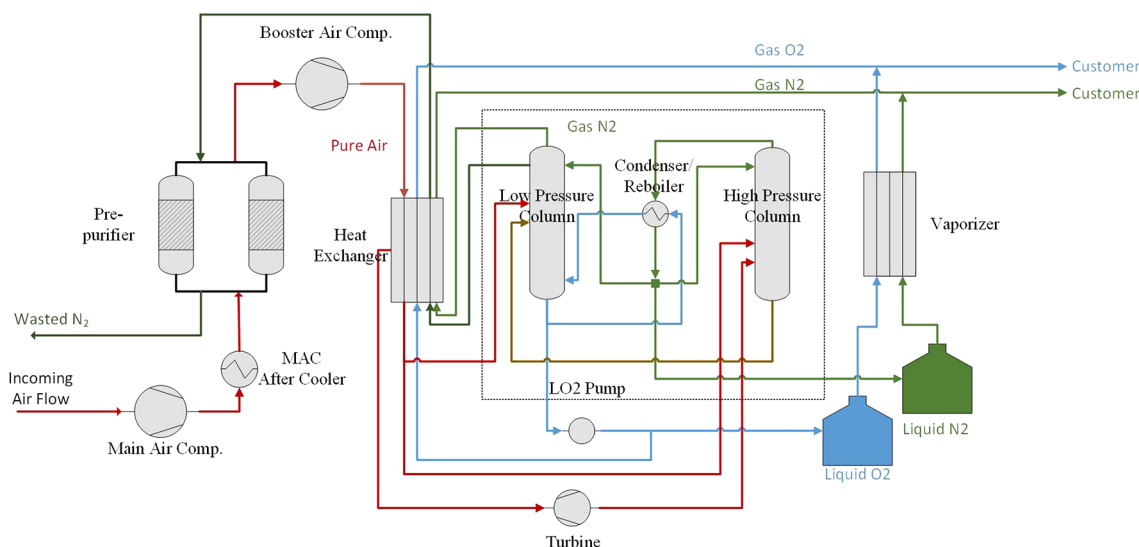


Figure 1. ASU with storage

paper models the stochastic process exactly, which allows it to be associated with outage penalty and be reflected in the objective function.

With the Markov Chain framework, the number of possible states propagate geometrically with the size of the process superstructure, which gives rise to the need of decomposition. However, while the exact modeling of outage penalties provides a closer representation of the trade-offs that decision makers face in reality, it also results in highly nonlinear and nonconvex functional relationships. Therefore, general decomposition methods such as Lagrangean decomposition and Benders decomposition cannot be readily applied. As an alternative perspective, the problem can be decomposed into a team game for which each processing stage is an individual agent. Plenty of works have incorporated game theory to address the competitive aspect of multiagent decision making as an improvement to the traditional holistic optimization. Lou et al.¹⁴ perform an emergy analysis of an industrial ecosystem that treats the member entities, such as plants, as players. Castillo et al.¹⁵ propose a concordant decision-making model based on game theory for liquified natural gas (LNG) processes. Zamarripa et al.¹⁶ address both the cooperative and competitive properties of supply chain planning problems with a game theory approach. On the other hand, a team game¹⁷ considers the cooperative aspect, in which the inputs of one agent include the actions of other agents. A major application of the concept is the area of distributed control^{18,19} for handling complex information structure. As presented in the following sections, the problem of concern has a similar information structure. Therefore, in this work, we use the concept of team game to reach a local optimum starting from a good initial solution. It provides a novel perspective for the simultaneous optimal design of system redundancy and storage with respect to product availability, on top of the proposed general mixed integer formulation, in which the stochastic process inferences of equipment failures and repairs are modeled exactly.

Section 2 introduces the motivating example of an air separation unit (ASU) process and formally states the problem. In section 3, we develop the MILP model (RST) and show that its computational complexity increases drastically with problem size. Section 4 presents the results of directly solving the motivating example with the MILP model (RST) as well as

the computational difficulty of applying the model to a slightly larger example. In section 5, we introduce an iterative algorithm toward the team game Nash equilibrium, which is a necessary condition for the global optimum. We show that a good initial solution can be easily obtained for our problem, which guarantees the quality of the Nash equilibrium solution. Section 6 shows a series of examples for which the global optima are obtained by the game theoretic approach in much shorter time than directly solving the original MILP model (RST).

2. MOTIVATING EXAMPLE

Figure 1 shows a typical air separation unit that has two liquid product storage tanks for LO₂ and LN₂, respectively. The air first goes through the main air compressor and the after cooler to be pressurized, then the prepurifier to remove impurities such as CO₂. After that, the air is compressed again by the booster air compressor and cooled by the gas product of nitrogen and liquid product of oxygen. Usually, about two-thirds of the air will be streamered out and expanded by a gas turbine before being fed into the high pressure column. The rest of the air is cooled down to be the two-phase region in the heat exchanger, which is then split into two streams, and fed into the low pressure column and the high pressure column separately. It is worth mentioning that the liquid O₂ comes out from the bottom of the low pressure column. Therefore, a pump is needed to bring the stream out, while the liquid N₂ product comes out from the high pressure column and does not need to be pressurized.

We analyze the reliability aspect of air separation units that are built next to a designated customer and has to constantly supply gas piped through pipelines. When the ASU is down due to equipment failures, the liquid products can be vaporized to sustain pipeline supply. However, if the downtime is too long such that the tank runs out before the ASU is back online, an interruption of pipeline supply is unavoidable. Typically, the gas supplier and the customers agree that each occurrence of pipeline interruption leads to a fixed penalty charge to the supplier.

Figure 2 shows the system superstructure. The purpose of the model is to determine the selection of parallel units among a finite number of possibly distinct candidates, and the sizes of

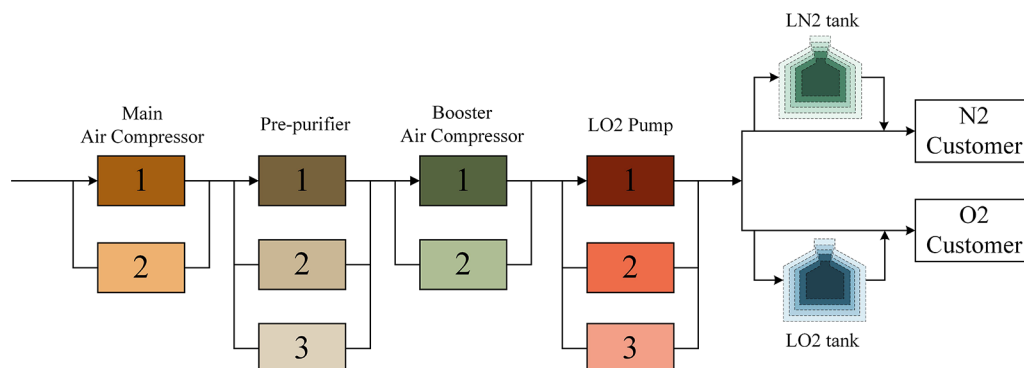


Figure 2. ASU with storage.

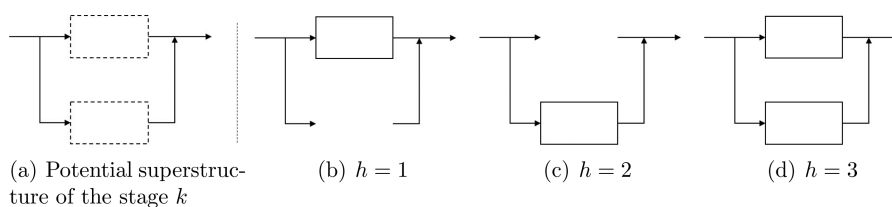


Figure 3. Design options of individual stages.

the two liquid product storage tanks, in order to achieve the minimum cost balancing capital investment and the interruption penalty from the pipeline customers.

3. MATHEMATICAL MODEL

The site is considered to have two parts, the air separation system with several stages in series, and the storage tanks. In section 3.1, the failure-repair process of the air separation system is modeled as a continuous-time Markov Chain depending on the system design, or redundancy selection. Next, in section 3.2, the impact of storage tank sizes on system availability is incorporated based on the relationships established in section 3.1.

3.1. Modeling the Processing System as a Continuous-Time Markov Chain. The processing stages, for example, main air compressor and prepurifier, are indexed with $k \in K$, whereas the individual potential redundancy designs in each stage are indexed with $h \in H_k$. Figure 3a shows a stage k with potentially two units, while Figures 3b–3d show the three potential redundancy designs (selections) indexed by h , based on the superstructure.

The binary variable $z_{k,h}$ indicates which design h is selected for stage k . Equation 1 requires that only one potential design be selected for each stage k . The investment cost of the design h in stage k is $\hat{C}_{k,h}^U$. Equation 2 calculates the investment cost of processing units depending on the design selection.

$$\sum_{h \in H_k} z_{k,h} = 1, \quad \forall k \in K \quad (1)$$

$$C^U = \sum_{k \in K} \sum_{h \in H_k} z_{k,h} \hat{C}_{k,h}^U \quad (2)$$

On the basis of the assumption that the time to failure and time to repair of single units follow exponential distributions, we model the failure–repair process of the processing stage k with design h as a continuous-time Markov Chain, and establish the corresponding transition matrix $\mathbf{W}_{k,h}$, from which two key quantities can be obtained: $\pi_{k,h}$ is the probability distribution

vector over the state space $S_{k,h}$ calculated from the corresponding transition matrix $\mathbf{W}_{k,h}$ with eq 3; $\sigma_{k,h}$ is the diagonal of $\mathbf{W}_{k,h}$ as stated in eq 4. By the definition of continuous-time Markov Chain, the residence time of any state follows an exponential distribution. The physical meaning of $\sigma_{k,h}(s)$ is that it represents the rate parameter of the exponential distribution of state $s \in S_{k,h}$. For a thorough explanation of Markov Chain please refer to the book by Sericola et al.²⁰ For the application of Markov Chain to redundancy systems please refer to our previous paper.¹² It is worth mentioning that we only consider independent failures that originated from within the individual units, and exclude those that are correlated and are due to external factors such as extreme weather, power fluctuation, etc.

$$\pi_{k,h}^T \cdot [\mathbf{W}_{k,h}, 1] = [0^T, 1] \quad (3)$$

$$\sigma_{k,h} = \text{diag}(\mathbf{W}_{k,h}) \quad (4)$$

With the stage level dynamics understood, we introduce the concept of system design, which is the combination of the individual stage designs. The set of all unique potential system designs is \bar{H} , and the elements are indexed by \bar{h} . Figure 4 shows the nine possible system designs of a system with two stages just like the one stage shown in Figure 3a.

Again, we model the failure–repair process of the system with design \bar{h} as a continuous-time Markov Chain with transition matrix $\bar{\mathbf{W}}_{\bar{h}}$. $\bar{\pi}_{\bar{h}}$ is the probability distribution vector over the state space $\bar{S}_{\bar{h}}$ of the system design \bar{h} calculated by eq 5. $\bar{\sigma}_{\bar{h}}$ is the diagonal of $\bar{\mathbf{W}}_{\bar{h}}$ and $\bar{\sigma}_{\bar{h}}(\bar{s})$ is the rate parameter of the exponential distribution dictating the residence time of state $\bar{s} \in \bar{S}_{\bar{h}}$.

$$\bar{\pi}_{\bar{h}}^T \cdot [\bar{\mathbf{W}}_{\bar{h}}, 1] = [0^T, 1] \quad (5)$$

$$\bar{\sigma}_{\bar{h}} = \text{diag}(\bar{\mathbf{W}}_{\bar{h}}) \quad (6)$$

In Appendix A in the Supporting Information we show that $\bar{\pi}_{\bar{h}}$ and $\bar{\sigma}_{\bar{h}}$ of system design \bar{h} can be calculated based on $\pi_{k,h}$ and $\sigma_{k,h}$ as shown in eqs 7–9, for which design $h(k, \bar{h})$ in stage k

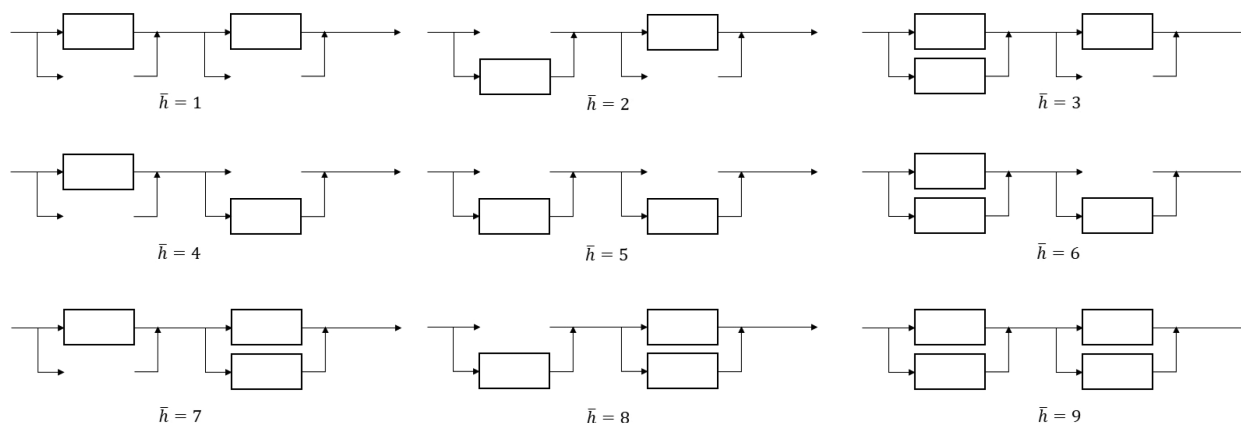


Figure 4. Potential system designs.

is part of system design \bar{h} . (The $\prod \otimes$ represents the Kronecker product. Please see Appendix A for illustration.)

$$\bar{\pi}_{\bar{h}} = \pi_{|K|,h(|K|,\bar{h})} \otimes \dots \otimes \pi_{2,h(2,\bar{h})} \otimes \pi_{1,h(1,\bar{h})} \quad (7)$$

$$\bar{\sigma}_{\bar{h}} = \sum_{k \in K} 1_{(g_k)} \otimes \sigma_{k,h(k,\bar{h})} \otimes 1_{(l_k)} \quad (8)$$

g_k is the product of the state space dimensions of all the stages after stage k in the stage sequence, while l_k is the product of the state space dimensions of all the stages before k .

$$g_k = \prod_{(l,h): l > k, (l,h,\bar{h}) \in HC} \|\pi_{l,h}\|_0, \\ l_k = \prod_{(l,h): l < k, (l,h,\bar{h}) \in HC} \|\pi_{l,h}\|_0 \quad (9)$$

As mentioned above, $rt_{\bar{h}}(\bar{s})$, the residence time of state $\bar{s} \in \bar{S}_{\bar{h}}$, follows the exponential distribution with rate parameter $\bar{\sigma}_{\bar{h}}(\bar{s})$.

$$P(rt_{\bar{h}}(\bar{s}) = t) = \bar{\sigma}_{\bar{h}}(\bar{s}) e^{-\bar{\sigma}_{\bar{h}}(\bar{s})t}, \quad \forall \bar{s} \in \bar{S}_{\bar{h}} \quad (10)$$

Finally, we introduce $f_{\bar{h}}(\bar{s})$, the frequency of encountering each state \bar{s} . It is equal to the long-term probability $\pi_{\bar{h}}(\bar{s})$ divided by the mean residence time $\overline{rt_{\bar{h}}}$, which is equal to the reciprocal of the rate parameter $\bar{\sigma}_{\bar{h}}(\bar{s})$.

$$f_{\bar{h}}(\bar{s}) = \frac{\pi_{\bar{h}}(\bar{s})}{\overline{rt_{\bar{h}}}} = \pi_{\bar{h}}(\bar{s}) \bar{\sigma}_{\bar{h}}(\bar{s}), \quad \forall \bar{s} \in \bar{S}_{\bar{h}} \quad (11)$$

3.2. System Availability. In this section, we discuss how to incorporate the size of the storage tank for evaluating the network availability. The main idea is to calculate the frequency of the incidents where the plant is down due to unplanned equipment failure, and the downtime is too long such that the liquid storage runs out. Furthermore, we account for the proportional penalization in the objective function, because the actual penalty being paid to the customers is directly proportional to the frequency of supply interruption.

In section 3.1, it is established that the air separation system follows a continuous-time Markov Chain with transition matrix $\bar{W}_{\bar{h}}$ for potential system design \bar{h} , which produces the stationary probability distribution $\bar{\pi}_{\bar{h}}(\bar{s})$ over the state space $(\bar{S}_{\bar{h}})$, the residence time rate parameter $\bar{\sigma}_{\bar{h}}(\bar{s})$, and the frequency of encountering each state $f_{\bar{h}}(\bar{s})$.

We first focus on the random variable of the volume of liquid products decreasing during failure state $\bar{s} \in \bar{S}_{\bar{h}}^f$. It is proportional to the consumption rates δ^{LO2} or δ^{LN2} , and the

random variable of the residence time $rt_{\bar{h}}(\bar{s})$. On the basis of the exponential distribution followed by $rt_{\bar{h}}(\bar{s})$ as shown in eq 10, they follow the exponential distribution with rate parameters $\frac{\bar{\sigma}_{\bar{h}}(\bar{s})}{\delta^{LO2}}$ or $\frac{\bar{\sigma}_{\bar{h}}(\bar{s})}{\delta^{LN2}}$.

$$P(V_{\bar{h}}^{LO2dec}(\bar{s}) = V) = \frac{\bar{\sigma}_{\bar{h}}(\bar{s})}{\delta^{LO2}} e^{-\bar{\sigma}_{\bar{h}}(\bar{s})/(\delta^{LO2}V)}, \quad \bar{s} \in \bar{S}_{\bar{h}}^f \quad (12)$$

$$P(V_{\bar{h}}^{LN2dec}(\bar{s}) = V) = \frac{\bar{\sigma}_{\bar{h}}(\bar{s})}{\delta^{LN2}} e^{-\bar{\sigma}_{\bar{h}}(\bar{s})/(\delta^{LN2}V)}, \quad \bar{s} \in \bar{S}_{\bar{h}}^f \quad (13)$$

We consider that there are a finite number of storage size options (k gallon) for both products, which we denote as V_n^{LO2} and V_n^{LN2} , with N^{LO2} and N^{LN2} as the index set of size options. x_n^{LO2} and x_n^{LN2} are binary variables that indicate the selection of tank sizes for LO2 (V_n^{LO2}) and LN2 (V_n^{LN2}), respectively. Equations 14 and 15 require that one and only one size be selected for each product.

$$\sum_{n \in N^{LO2}} x_n^{LO2} = 1 \quad (14)$$

$$\sum_{n \in N^{LN2}} x_n^{LN2} = 1 \quad (15)$$

C^T is the investment cost of storage tanks depending on their size selections.

$$C^T = \sum_{n \in N^{LO2}} x_n^{LO2} c_n^{LO2} + \sum_{n \in N^{LN2}} x_n^{LN2} c_n^{LN2} \quad (16)$$

on the basis of eqs 11–13, the outage frequencies of LO2 and LN2 for system design \bar{h} and tank size n for the respective products, $\widehat{fr}_{n,\bar{h}}^{LO2}$ and $\widehat{fr}_{n,\bar{h}}^{LN2}$, can be calculated with eqs 17 and 18.

$$\begin{aligned} \widehat{fr}_{n,\bar{h}}^{LO2} &= \sum_{\bar{s} \in \bar{S}_{\bar{h}}^f} f_{\bar{h}}(\bar{s}) \cdot P(V_{\bar{h}}^{LO2dec}(\bar{s}) \geq V_n^{LO2}) \\ &= \sum_{\bar{s} \in \bar{S}_{\bar{h}}^f} \bar{\pi}_{\bar{h}}(\bar{s}) \cdot \bar{\sigma}_{\bar{h}}(\bar{s}) \cdot e^{-V_n^{LO2}/(\delta^{LO2}\bar{\sigma}_{\bar{h}}(\bar{s}))} \end{aligned} \quad (17)$$

$$\begin{aligned} \widehat{fr}_{n,\bar{h}}^{LN2} &= \sum_{\bar{s} \in \bar{S}_{\bar{h}}^f} f_{\bar{h}}(\bar{s}) \cdot P(V_{\bar{h}}^{LN2dec}(\bar{s}) \geq V_n^{LN2}) \\ &= \sum_{\bar{s} \in \bar{S}_{\bar{h}}^f} \bar{\pi}_{\bar{h}}(\bar{s}) \cdot \bar{\sigma}_{\bar{h}}(\bar{s}) \cdot e^{-V_n^{LN2}/(\delta^{LN2}\bar{\sigma}_{\bar{h}}(\bar{s}))} \end{aligned} \quad (18)$$

The following constraints and disjunctions represent the relationship between the redundancy and storage selection and the expected outage penalties PN . In eq 19, T is the total time horizon being considered. $\text{penalty}^{\text{LO2}}$ and $\text{penalty}^{\text{LN2}}$ are the amount of penalty charge per outage based on the contracts between the gas supplier and the customer. Disjunctions (20–21) enforce that $f_{n,\bar{h}}^{\text{LO2}}$ and $f_{n,\bar{h}}^{\text{LN2}}$ are equal to the parameters $\widehat{f}_{n,\bar{h}}^{\text{LO2}}$ and $\widehat{f}_{n,\bar{h}}^{\text{LN2}}$ where the corresponding stage designs and storage tank sizes are selected.

$$PN = T(\text{penalty}^{\text{LO2}} \cdot f_{n,\bar{h}}^{\text{LO2}} + \text{penalty}^{\text{LN2}} \cdot f_{n,\bar{h}}^{\text{LN2}}) \quad (19)$$

$$\bigvee_{n \in N^{\text{LO2}}, \bar{h} \in \bar{H}} \left[X_n^{\text{LO2}} \wedge \left(\bigwedge_{k \in K} Z_{k,h(k,\bar{h})} \right) \right] \quad f_{n,\bar{h}}^{\text{LO2}} = \widehat{f}_{n,\bar{h}}^{\text{LO2}} \quad (20)$$

$$\bigvee_{n \in N^{\text{LN2}}, \bar{h} \in \bar{H}} \left[X_n^{\text{LN2}} \wedge \left(\bigwedge_{k \in K} Z_{k,h(k,\bar{h})} \right) \right] \quad f_{n,\bar{h}}^{\text{LN2}} = \widehat{f}_{n,\bar{h}}^{\text{LN2}} \quad (21)$$

We define the continuous variables $f_{n,\bar{h}}^{\text{LO2}}$ and $f_{n,\bar{h}}^{\text{LN2}}$ for the hull reformulation²¹ given by eqs 22–28.

$$PN = T(\text{penalty}^{\text{LO2}} \cdot \sum_{n \in N^{\text{LO2}}} \sum_{\bar{h} \in \bar{H}} f_{n,\bar{h}}^{\text{LO2}} + \text{penalty}^{\text{LN2}} \cdot \sum_{n \in N^{\text{LN2}}} \sum_{\bar{h} \in \bar{H}} f_{n,\bar{h}}^{\text{LN2}}) \quad (22)$$

$$f_{n,\bar{h}}^{\text{LO2}} \leq x_n^{\text{LO2}} \widehat{f}_{n,\bar{h}}^{\text{LO2}}, \quad \forall k \in K, n \in N^{\text{LO2}} \quad (23)$$

$$f_{n,\bar{h}}^{\text{LO2}} \leq z_{k,h(k,\bar{h})} \widehat{f}_{n,\bar{h}}^{\text{LO2}}, \quad \forall n \in N^{\text{LO2}}, \bar{h} \in \bar{H} \quad (24)$$

$$f_{n,\bar{h}}^{\text{LO2}} \geq \left(\sum_{k \in K} z_{k,h(k,\bar{h})} - |K| + x_n^{\text{LO2}} \right) \widehat{f}_{n,\bar{h}}^{\text{LO2}}, \quad \forall n \in N^{\text{LO2}}, \bar{h} \in \bar{H} \quad (25)$$

$$f_{n,\bar{h}}^{\text{LN2}} \leq x_n^{\text{LN2}} \widehat{f}_{n,\bar{h}}^{\text{LN2}}, \quad \forall k \in K, n \in N^{\text{LN2}} \quad (26)$$

$$f_{n,\bar{h}}^{\text{LN2}} \leq z_{k,h(k,\bar{h})} \widehat{f}_{n,\bar{h}}^{\text{LN2}}, \quad \forall n \in N^{\text{LN2}}, \bar{h} \in \bar{H} \quad (27)$$

$$f_{n,\bar{h}}^{\text{LN2}} \geq \left(\sum_{k \in K} z_{k,h(k,\bar{h})} - |K| + x_n^{\text{LN2}} \right) \widehat{f}_{n,\bar{h}}^{\text{LN2}}, \quad \forall n \in N^{\text{LN2}}, \bar{h} \in \bar{H} \quad (28)$$

The objective function in eq 29 minimizes the total investment cost and the expected penalty.

$$\min_{z_{k,h}, x_n^{\text{LO2}}, x_n^{\text{LN2}}} C^U + C^T + PN \quad (29)$$

To summarize, the model (RST) is an MILP that minimizes the total cost eq 29 under the constraints 1, 2, 14–16, and 22–28. A hidden bottleneck in the model (RST) is the computation of the exponentially many numbers of parameters $\widehat{f}_{n,\bar{h}}^{\text{LO2}}$ and $\widehat{f}_{n,\bar{h}}^{\text{LN2}}$ (based on eqs 3, 4, 7, 8, 17, and 18) with respect to the size of the design superstructure. To illustrate, we consider a small example for the potential designs in Figure 4 with three possible sizes for LO2 and LN2 storage tanks. To determine the 18 parameters for each of the two groups

$\widehat{f}_{n,\bar{h}}^{\text{LO2}}$ and $\widehat{f}_{n,\bar{h}}^{\text{LN2}}$, we have to solve for the probability vectors of nine different state spaces of dimensions 4–25, before carrying out the calculation in eqs 17 and 18 for the different tank sizes. Table 1 shows how the computational workload

Table 1. Required Computations for Different Superstructures

		small	medium	large
system statistics	number of stages	2	3	4
	number of potential units per stage	2	3	4
	number of potential tank sizes	3	4	5
computational tasks	number of state space	9	343	50 625
	max dimension of state space	24	21 870	141 087 744
	number of parameters	36	2744	506 250

increases drastically as the superstructures grow in size. Specifically, to calculate a single parameter $\widehat{f}_{n,\bar{h}}^{\text{LO2}}$, we need to find out the corresponding state space of potential design \bar{h} , and calculate $\bar{\pi}_h^T$ and $\bar{\sigma}_h^T$, the complexity of which depends on the dimension of the state space. Therefore, not only does the number of parameters but also the computational effort needed to determine each single parameter increase with the size of the superstructure.

4. SOLUTION OF THE MOTIVATING PROBLEM AND RESTRICTIONS OF THE MODEL (RST)

In this section, we present more details about the example shown in section 2, including the assigned parameters and corresponding solutions.

Table 2 shows the failure modes considered for each processing stage. The mean time between failures (MTBF) and the mean time to repair (MTTR) are known for each failure mode. The failure rates and repair rates can be obtained by taking their respective reciprocals. The transition matrices are obtained as discussed in our previous work.¹² The mean time between failures (MTBF) ranges from 5 to 25 years. Mean time to repair (MTTR) ranges from 8 to 1080 h.

The capital cost of each unit ranges from \$140k to \$1250k. Table 3 shows the penalty rates and pipeline flow rates used in the model.

Table 4 shows the tank size options and corresponding costs for LO2 and LN2 storage. LO2 tanks are generally more expensive.

The MILP model (RST) has 8828 equations and 2800 variables, with 269 of them being binary variables. The parameters are calculated in Python and the model is solved with CPLEX 12.8.0.0 in Pyomo. The total time to calculate the parameters and solving the model is 8.0 s, where 5.3 s is used for parameter calculation, and 2.7 s is used for solving the MILP. The optimal design is shown in Figure 5. The expected frequency of LO2 outage is 0.005732 in the 10 year horizon, which incurs a penalty of \$11,464. The expected frequency of LN2 outage is 0.006 in the 10 year horizon, which incurs a \$12,291 penalty.

As mentioned in section 3.2, it becomes more challenging to enumerate all possible combinations of (n, \bar{h}) and calculate $\widehat{f}_{n,\bar{h}}^{\text{LO2}}$ and $\widehat{f}_{n,\bar{h}}^{\text{LN2}}$ as the number of stages and potential units increase. In fact, for a larger problem where there is one more potential unit for each of first three stages, the computational

Table 2. Failure Modes of Each Processing Stage

stage	failure mode	stage	failure mode	stage	failure mode	stage	failure mode
main air compressor	FMC1	prepuriifier	FMPF1	booster air compressor	FMC1	LO ₂ Pump	FMP1
	FMC2				FMC2		
	FMC3				FMC3		
	FMC4				FMC4		
	FMC5				FMC5		
	FMC6				FMC6		

Table 3. Profitability parameters

penalty ^{LO2} (k\$ per outage)	penalty ^{LN2} (k\$ per outage)	δ^{LO2} (k gallon per day)	δ^{LN2} (k gallon per day)
2000	2000	48	60

Table 4. Tank Size Options and Costs

size options (k gallon)	100	400	700	1000	1500
price for LO ₂ (k\$)	55	237	427	621	951
price for LN ₂ (k\$)	50	215	388	565	864

time of (RST) is 2053.5 s, out of which 1697.2 s are used for the parameter calculation. Moreover, for the example presented above, the total number of complicating parameters is 2520, and the maximum number of scenarios to consider for each parameter is 4536, while for the larger problem, those numbers become 37 730 and 14 117 880, respectively.

5. ITERATIVE GAME-THEORETIC ALGORITHM

To overcome the bottleneck of the exponential parameter computational time, we propose an algorithm that starts with approximate and incomplete parameters, and iteratively adjusts the parameter pool as appropriate while the models are being solved, which allows us to achieve a Nash equilibrium, or a "person-by-person" optimum, among the stages.

The analogy of the problem being addressed to a team game is as follows. Each stage has its own contribution to the system costs by adding capital costs as well as the penalties from unplanned downtimes. The condition of each stage (selection of parallel units) impacts its contribution to the penalties, just like each player is largely responsible for his/her contribution to the team performance. However, they play as a team in the sense that the individual contributions to the penalties (if a failure is due to a specific stage, the resulting downtime is considered its contribution) can be slightly affected by the conditions of the other stages. To be more specific, a failure of one stage means the failure of the entire system. Therefore, the extents of reliability and investment levels have to be balanced

among the stages. For example, it might not be favorable to have one stage being much more reliable than the others, even if it is the best solution for itself standing alone. It is because the other stages are more likely to fail before it does and cover its failures, and therefore, the marginal decrease in its contribution to the system failure from increasing reliability investment is less than the case where this stage forms a one-stage-system.

The problem can be viewed as a team game²² in which the players are the $|K|$ stages. Inequality 30 provides the definition of Nash equilibrium in a team game: Given the strategy of all other players $\rho' \in \mathcal{P}$ (the design selection of all other stages $l \in K, l \neq k$), there is no other strategy $\xi'_\rho \in \Xi_\rho$ (potential design h) for any single player $\rho \in \mathcal{P}$ (single stage k) that could give better utility C (lower objective function values):

$$C(\xi_1^*, \dots, \xi_\rho^*, \xi_{|\mathcal{P}|}^*) \leq C(\xi_1^*, \dots, \xi'_\rho, \xi_{|\mathcal{P}|}^*), \quad \forall \rho \in \mathcal{P}, \xi'_\rho \in \Xi_\rho \quad (30)$$

Before going into the details of the proposed algorithm, we will graphically illustrate its significance of reducing parameter calculations by considering a small system as shown in Figure 6. The three rectangles grouping three dots of different sizes represent three stages each with three design options. The original formulation (RST) requires calculating the parameters for each valid combination (27 triangles that have one vertex in each rectangle).

However, in Figures 7, 8, and 9, which represent the steps in the proposed algorithm, the connecting segments can be much sparser, since much fewer parameters need to be calculated following the proposed algorithm pursuing a Nash equilibrium.

First, we optimize the stage-wise design selection with approximated impact from other stages. The approximation effort is presented in detail in section 5.1. The triangle in Figure 7 stands for the selected designs for individual stages.

An equilibrium test is then performed to this group of designs to determine whether any single stage has a superior deviation (deviation that generates lower total cost) given that the rest of the stages stick to the current plan. As shown in

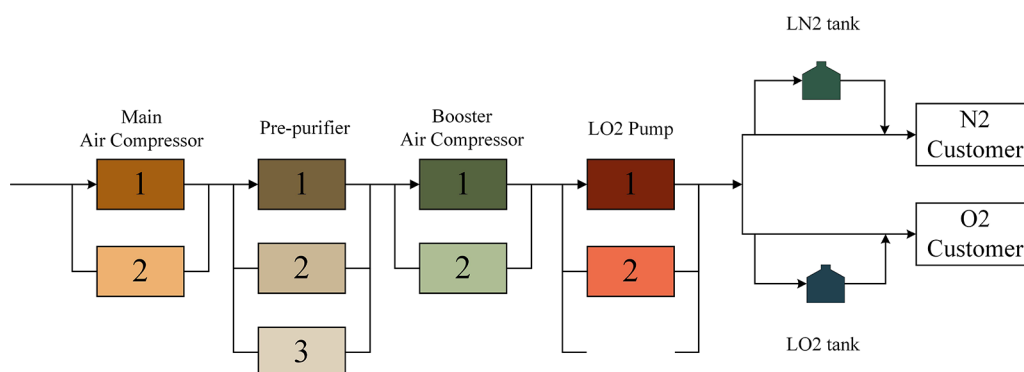


Figure 5. ASU with storage

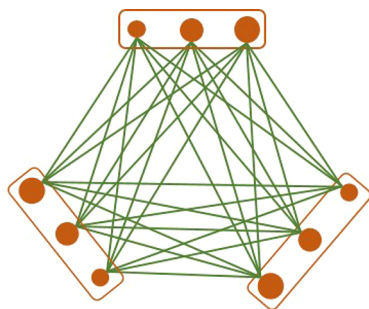


Figure 6. The original way has to calculate everything in advance.

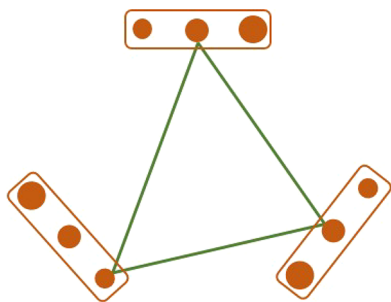


Figure 7. Initial approximated optimum.

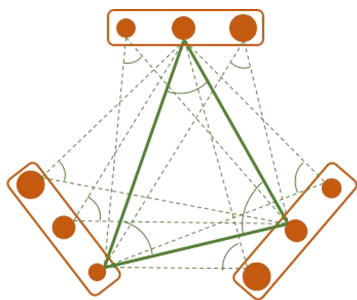


Figure 8. Equilibrium checking.

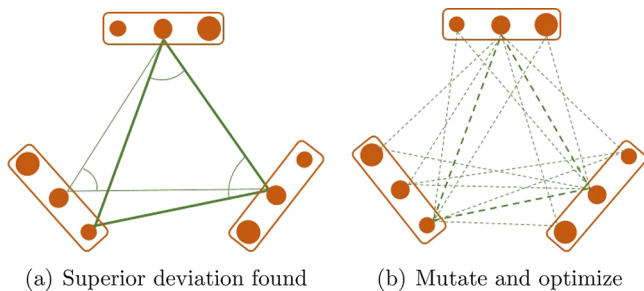


Figure 9. When current solution does not satisfy equilibrium condition.

Figure 8, the solid triangle represents the group of designs undergoing the equilibrium test following from Figure 7. In each rectangle group of dots, the two dots supporting dashed vertices represent the deviations being examined, which oppose the solid sides, as the deviations are taken with the rest of the stages unchanged and unaffected.

If there is no superior deviation to the current plan, then the current design is a Nash equilibrium point. If otherwise, for example, in Figure 9a, the dot supporting the vertex with thinner sides is a superior deviation for the stage to which it belongs, we go to the next step shown in Figure 9b. In this step, we add all the design selections (the six triangles with two thin dashed sides) that can be obtained by alternating the current selection,

to the pool of potential system designs, while removing the current selection itself (the triangle with all thick dashed sides).

It is worth noting the difference between the six vertices supported by dashed sides in Figure 8 and the six dashed triangles in Figure 9b: The dashed vertices are fictional deviations that affect the performance of the stage itself but not the outputs of the other stages (players), while a dashed triangle means that all the stages (players) take on a new set of designs (strategies) together, and the impacts on the outputs of each other are accounted for.

The optimal design in the current pool will again go back through the equilibrium check stage like in Figure 8, until a Nash equilibrium point is found.

We can see that the new algorithm reduces a large part of the parameter calculations compared to the original model, especially when the number of stages and potential designs increases. It is worth mentioning that an overall optimum implies a Nash equilibrium, but the reverse is not true. Therefore, the Nash equilibrium that we pursue in this method can be considered a suboptimal solution or a local optimum.

In section 5.1, we first show how the complicating parameters, $\widehat{fr}_{n,\bar{h}}^{LO2}$ and $\widehat{fr}_{n,\bar{h}}^{LN2}$ in eqs 17 and 18, can be closely approximated with much simpler expressions. Next we present a rearrangement of the terms that breaks down the summations over system failure states according to the exact stage where the failure is happening, which reveals a linear impact of the other stages on the contribution to liquid-overconsumption of each individual stages. In section 5.2, we present a step-by-step description of the algorithm.

5.1. Approximation and Rearrangement of Key Parameter Expression. In this section, we propose an approximate expression and a rearranged expression of the key parameters, the outage frequencies of the liquid products (LO2 and LN2) output, $\widehat{fr}_{n,\bar{h}}^{LO2}$ and $\widehat{fr}_{n,\bar{h}}^{LN2}$, which is in preparation for the iterative solution scheme explained in the next section. In the following context, we only focus on LO2 and omit the similar term rearrangements and approximations that are also done for LN2 to avoid unnecessary repetition.

First, we rearrange the expression in eq 31 with the following notations:

For each stage k , we let \tilde{h} index the potential designs of the complementary subsystem obtained by subtracting stage k from the entire system. For example, as shown in Figure 10, in the perspective of stage 1, the complementary subsystem refers to the system consisting of stages 2 and 3. Similarly to $h(k, \bar{h})$, we define $\tilde{h}(k, \bar{h})$ as the index of the potential complementary design of stage k that is part of system design \bar{h} .

In the original expression shown in eq 17, the system failure scenarios are directly enumerated, whereas in eq 31, the summation over the system failure scenarios breaks down by the stage k that causes the failure (see Appendix B for detailed derivation).

$$\widehat{fr}_{n,\bar{h}}^{LO2} = \sum_{k \in K} \sum_{s \in S_{k,h(k,\bar{h})}^f} \left[\pi_{k,h(k,\bar{h})}(s) \cdot e^{-(V_n^{LO2}/\delta^{LO2})\sigma_{k,h(k,\bar{h})}(s)} \cdot \sum_{\tilde{s} \in \tilde{S}_{k,\tilde{h}(k,\bar{h})}} \tilde{\pi}_{k,\tilde{h}(k,\bar{h})}(\tilde{s}) \cdot e^{-(V_n^{LO2}/\delta^{LO2})\tilde{\sigma}_{k,\tilde{h}(k,\bar{h})}(\tilde{s})} \cdot \left(1 + \frac{\tilde{\sigma}_{k,\tilde{h}(k,\bar{h})}(\tilde{s})}{\sigma_{k,h(k,\bar{h})}(s)} \right) \right] \quad (31)$$

We denote the stage-wise component of $\widehat{fr}_{n,\bar{h}}^{LO2}$ as $\widehat{FR}_{n,k,h(k,\bar{h})}^{LO2}$:

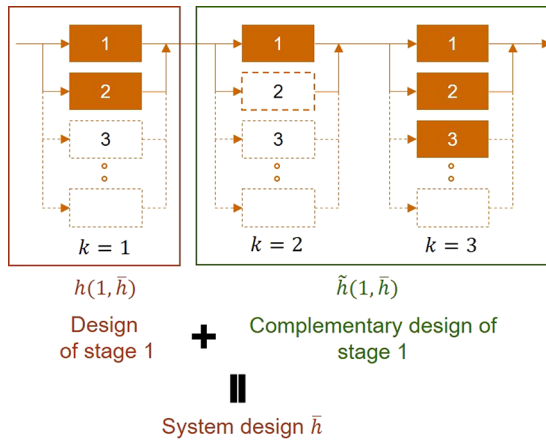


Figure 10. From the perspective of stage 1.

$$\widehat{f}_{n,\bar{h}}^{\text{LO2}} = \sum_{k \in K} \widehat{FR}_{n,k,h(k,\bar{h})}^{\text{LO2}} \quad (32)$$

After rearrangement, $\widehat{FR}_{n,k,h(k,\bar{h})}^{\text{LO2}}$ has the exact expression as shown in eq 33.

$$\widehat{FR}_{n,k,h(k,\bar{h})}^{\text{LO2}} = \Theta_{n,k,h(k,\bar{h})}^{\text{LO2}} \cdot \Phi_{n,k,h(k,\bar{h})}^{\text{LO2}} + \Phi_{n,k,h(k,\bar{h})}^{\text{LO2}} \cdot \Theta_{n,k,h(k,\bar{h})}^{\text{LO2}} \quad (33)$$

where

$$\Theta_{n,k,h}^{\text{LO2}} = \sum_{s \in S_{k,h}^f} \pi_{k,h}(s) \cdot e^{-(V_n^{\text{LO2}}/\delta^{\text{LO2}})\sigma_{k,h}(s)} \cdot \sigma_{k,h}(s) \quad (34)$$

$$\Phi_{n,k,h}^{\text{LO2}} = \sum_{\tilde{s} \in \tilde{S}_{k,h}} \tilde{\pi}_{k,h}(\tilde{s}) \cdot e^{-(V_n^{\text{LO2}}/\delta^{\text{LO2}})\tilde{\sigma}_{k,h}(\tilde{s})} \quad (35)$$

$$\Phi_{n,k,h}^{\text{LO2}} = \sum_{s \in S_{k,h}^f} \pi_{k,h}(s) \cdot e^{-(V_n^{\text{LO2}}/\delta^{\text{LO2}})\sigma_{k,h}(s)} \quad (36)$$

$$\Theta_{n,k,h}^{\text{LO2}} = \sum_{\tilde{s} \in \tilde{S}_{k,h}} \tilde{\sigma}_{k,h}(\tilde{s}) \cdot e^{-(V_n^{\text{LO2}}/\delta^{\text{LO2}})\sigma_{k,h}(\tilde{s})} \cdot \tilde{\sigma}_{k,h}(\tilde{s}) \quad (37)$$

Therefore, instead of enumerating by system designs \bar{h} , we focus on one stage k at a time and enumerate by its potential designs $h \in H_k$ and the potential complementary designs $\tilde{h} \in \tilde{H}_k$, and define $\widehat{FR}_{n,k,h,\tilde{h}}^{\text{LO2}}$:

$$\widehat{FR}_{n,k,h,\tilde{h}}^{\text{LO2}} = (\Theta_{n,k,h}^{\text{LO2}} \cdot \Phi_{n,k,\tilde{h}}^{\text{LO2}} + \Phi_{n,k,h}^{\text{LO2}} \cdot \Theta_{n,k,\tilde{h}}^{\text{LO2}}) \quad (38)$$

It is also shown in Appendix B that following from eq 31, eqs 39 and 40 are based on the assumption that the failure rates are of the order of 10^{-4} , and the repair rates are of the order of 10^{-2} .

$$\widehat{f}_{n,\bar{h}}^{\text{LO2}} = \sum_{k \in K} \sum_{s \in S_{k,h(k,\bar{h})}^f} \pi_{k,h(k,\bar{h})}(s) \cdot e^{-(V_n^{\text{LO2}}/\delta^{\text{LO2}})\sigma_{k,h(k,\bar{h})}(s)} \cdot \sigma_{k,h(k,\bar{h})}(s) \cdot (1 \pm O(10^{-2})) \quad (39)$$

$$\widehat{FR}_{n,k,h(k,\bar{h})}^{\text{LO2}} = \Theta_{n,k,h(k,\bar{h})}^{\text{LO2}} \cdot (1 \pm O(10^{-2})) \quad (40)$$

Therefore, as shown in eq 40, we can use the sum of the outage frequencies of each stage alone $\sum_{k \in K} \Theta_{n,k,h(k,\bar{h})}^{\text{LO2}}$ to approximate the system outage frequency $\widehat{f}_{n,\bar{h}}^{\text{LO2}}$.

As mentioned at the beginning of this section, similar term rearrangements and approximations are done for LN2, but are omitted to avoid unnecessary repetition. To summarize, it is

proposed to rewrite the original expression shown in eq 17 as 41 and approximate $\widehat{f}_{n,\bar{h}}^{\text{LO2}}$ with eq 42.

$$\widehat{f}_{n,\bar{h}}^{\text{LO2}} = \sum_{k \in K} \widehat{FR}_{n,k,h(k,\bar{h}),\tilde{h}(k,\bar{h})}^{\text{LO2}} \quad (41)$$

$$\widehat{f}_{n,\bar{h}}^{\text{LO2}} \approx \sum_{k \in K} \Theta_{n,k,h(k,\bar{h})}^{\text{LO2}} \quad (42)$$

5.2. Algorithm Description. In this section, we provide a step-by-step description of the algorithm with reference to the above mathematical details and qualitative explanations. Briefly speaking, the algorithm takes advantage of the fact that the individual contribution from each stage is under limited impact of the other stages, and obtains a good initial solution ignoring these interrelations. It then iterates between two tasks, checking the equilibrium condition of a current design, and until confirming an equilibrium point, expanding the solution pool with all the mutations of unsuccessful previous designs to find a better design.

Step 1. Calculate $\Theta_{n,k,h}^{\text{LO2}}$ and $\Phi_{n,k,h}^{\text{LO2}}$ (first defined in 34 and 36). Solve the MILP independent design selection problem (IDS) defined below, featuring the approximated parameters $\Theta_{n,k,h}^{\text{LO2}}$ and $\Theta_{n,k,h}^{\text{LN2}}$. On the basis of the relationship shown in eq 42, the MILP (IDS) is an approximation of (RST).

$$(\text{IDS}): \min_{z_{k,h}, x_n^{\text{LO2}}, x_n^{\text{LN2}}} C^U + C^T + PN$$

$$\text{s.t.} \sum_{h \in H_k} z_{k,h} \geq 1, \quad \forall k \in K$$

$$C^U = \sum_{k \in K} \sum_{h \in H_k} z_{k,h} \hat{C}_{k,h}^U$$

$$\sum_{n \in N^{\text{LO2}}} x_n^{\text{LO2}} \geq 1$$

$$\sum_{n \in N^{\text{LN2}}} x_n^{\text{LN2}} \geq 1$$

$$C^T = \sum_{n \in N^{\text{LO2}}} x_n^{\text{LO2}} c_n^{\text{LO2}} + \sum_{n \in N^{\text{LN2}}} x_n^{\text{LN2}} c_n^{\text{LN2}}$$

$$f_{n,k,h}^{\text{LO2}} \leq z_{k,h} \Theta_{n,k,h}^{\text{LO2}}, \quad \forall n \in N^{\text{LO2}}, k \in K, h \in H_k$$

$$f_{n,k,h}^{\text{LO2}} \leq x_n^{\text{LO2}} \Theta_{n,k,h}^{\text{LO2}}, \quad \forall n \in N^{\text{LO2}}, k \in K, h \in H_k$$

$$f_{n,k,h}^{\text{LO2}} \geq (z_{k,h} + x_n^{\text{LO2}} - 1) \Theta_{n,k,h}^{\text{LO2}}, \quad \forall n \in N^{\text{LO2}}, k \in K, h \in H_k$$

$$f_{n,k,h}^{\text{LN2}} \leq z_{k,h} \Theta_{n,k,h}^{\text{LN2}}, \quad \forall n \in N^{\text{LN2}}, k \in K, h \in H_k$$

$$f_{n,k,h}^{\text{LN2}} \leq x_n^{\text{LN2}} \Theta_{n,k,h}^{\text{LN2}}, \quad \forall n \in N^{\text{LN2}}, k \in K, h \in H_k$$

$$f_{n,k,h}^{\text{LN2}} \geq (z_{k,h} + x_n^{\text{LN2}} - 1) \Theta_{n,k,h}^{\text{LN2}}, \quad \forall n \in N^{\text{LN2}}, k \in K, h \in H_k$$

$$PN = T(\text{penalty}^{\text{LO2}} \cdot \sum_{n \in N^{\text{LO2}}} \sum_{k \in K} \sum_{h \in H_k} f_{n,k,h}^{\text{LO2}} + \text{penalty}^{\text{LN2}} \cdot \sum_{n \in N^{\text{LN2}}} \sum_{k \in K} \sum_{h \in H_k} f_{n,k,h}^{\text{LN2}})$$

$$C^U, C^T, PN \geq 0$$

$$z_{k,h} \in \{0,1\}, \quad \forall k \in K, h \in H_k$$

$$x_n^{\text{LO2}} \in \{0,1\}, \quad \forall n \in N^{\text{LO2}}$$

$$x_n^{\text{LN2}} \in \{0,1\}, \quad \forall n \in N^{\text{LN2}} \quad (43)$$

The optimal solution of (IDS) indicates the potential design to select for each stage k : $\{h_{(0)}^*(k), k \in K\}$ (refer to the dots supporting the triangle in Figure 7). Appendix C shows an estimated bound of the gap between the global optimum and the objective function resulting from k : $\{h_{(0)}^*(k), k \in K\}$ and the best storage selections that go with it.

Step 2. For each stage $k \in K$, solve the MILP equilibrium checking problems (ECP) $_{(m,k)}$. m starts counting at $m = 1$. The model formulation is shown in eq 45. To do that, we have to first calculate the parameters $\widehat{FR}_{n,k,h,m}^{LO2}$ based on $\{h_{(m-1)}^*(k), k \in K\}$ as shown in ref 44 (refer to the dashed vertices in Figure 8).

$$\widehat{FR}_{n,k,h,m}^{LO2} = \Theta_{n,k,h}^{LO2} \cdot \Phi_{n,k,h}^{LO2} + \Phi_{n,k,h}^{LO2} \cdot \Theta_{n,k,h}^{LO2} \quad (44)$$

$$(ECP)_{(m,k)}: \min_{z_{k,h}, x_n^{LO2}, x_n^{LN2}} C^U + C^T + PN$$

$$\text{s.t. } \sum_{h \in H_k} z_{k,h} \geq 1, \quad \forall k \in K$$

$$C^U = \sum_{h \in H_k} z_{k,h} \hat{C}_{k,h}^U + \sum_{l \in K \setminus \{k\}} \sum_{h \in H_l} \hat{C}_{l,h}^{U*}$$

$$\sum_{n \in N^{LO2}} x_n^{LO2} \geq 1, \quad \sum_{n \in N^{LN2}} x_n^{LN2} \geq 1$$

$$C^T = \sum_{n \in N^{LO2}} x_n^{LO2} c_n^{LO2} + \sum_{n \in N^{LN2}} x_n^{LN2} c_n^{LN2}$$

$$f_{n,k,h,m}^{LO2} \leq z_{k,h} \widehat{FR}_{n,k,h,m}^{LO2}, \quad \forall n \in N^{LO2}, h \in H_k$$

$$f_{n,k,h,m}^{LO2} \leq x_n^{LO2} \widehat{FR}_{n,k,h,m}^{LO2}, \quad \forall n \in N^{LO2}, h \in H_k$$

$$f_{n,k,h,m}^{LO2} \geq (z_{k,h} + x_n^{LO2} - 1) \widehat{FR}_{n,k,h,m}^{LO2}, \quad \forall n \in N^{LO2}, h \in H_k$$

$$f_{n,l,h,m}^{LO2} = 0, \quad \forall n \in N^{LO2}, l \in K \setminus \{k\}, h \in H_l \setminus \{h_{(m-1)}^*(k)\}$$

$$f_{n,l,h,m}^{LO2} = x_n^{LO2} \widehat{FR}_{n,l,h,m}^{LO2}, \quad \forall n \in N^{LO2}, l \in K \setminus \{k\}, h \in H_l \setminus \{h_{(m-1)}^*(k)\}$$

$$f_{n,k,h,m}^{LN2} \leq z_{k,h} \widehat{FR}_{n,k,h,m}^{LN2}, \quad \forall n \in N^{LN2}, h \in H_k$$

$$f_{n,k,h,m}^{LN2} \leq x_n^{LN2} \widehat{FR}_{n,k,h,m}^{LN2}, \quad \forall n \in N^{LN2}, h \in H_k$$

$$f_{n,k,h,m}^{LN2} \geq (z_{k,h} + x_n^{LN2} - 1) \widehat{FR}_{n,k,h,m}^{LN2}, \quad \forall n \in N^{LN2}, h \in H_k$$

$$f_{n,l,h,m}^{LN2} = 0, \quad \forall n \in N^{LN2}, l \in K \setminus \{k\}, h \in H_l \setminus \{h_{(m-1)}^*(k)\}$$

$$f_{n,l,h,m}^{LN2} = x_n^{LN2} \widehat{FR}_{n,l,h,m}^{LN2}, \quad \forall n \in N^{LN2}, l \in K \setminus \{k\}, h \in H_l \setminus \{h_{(m-1)}^*(k)\}$$

$$PN = T(\text{penalty}^{LO2} \cdot \sum_{n \in N^{LO2}} \sum_{k \in K} \sum_{h \in H_k} f_{n,k,h,m}^{LO2} + \text{penalty}^{LN2} \cdot \sum_{n \in N^{LN2}} \sum_{k \in K} \sum_{h \in H_k} f_{n,k,h,m}^{LN2})$$

$$\sum_{n \in N^{LN2}} \sum_{k \in K} \sum_{h \in H_k} f_{n,k,h,m}^{LN2}$$

$$c^U, c^T, PN \geq 0$$

$$z_{k,h} \in \{0,1\}, \quad \forall k \in K, h \in H_k$$

$$x_n^{LO2} \in \{0,1\}, \quad \forall n \in N^{LO2}$$

$$x_n^{LN2} \in \{0,1\}, \quad \forall n \in N^{LN2}$$

(45)

The optimal solutions of the models (ECP) $_{(m,k)}$, $k \in K$ indicate a group of design selection for each stage $\{h_{(m)}^*(k), k \in K\}$. If for every stage k , the complementary design indexed by (k, m) that was brought into (ECP) $_{(m)}$ consists of the selected

designs of the other stages, we say that a Nash equilibrium is found among the stages, and the iteration stops. Otherwise, go to step 3. The latter case is illustrated by Figure 9a.

Step 3. Solve the MILP partial design selection model (PDS) $_{(m)}$ given by eq 49, which corresponds to the alternate and optimize the step shown in Figure 9b. The model is called partial because only an incomplete pool of possible designs are being evaluated.

The associated parameters $\hat{C}_{\bar{h}}^U$, $\widehat{fr}_{n,\bar{h}}^{LO2}$ and $\widehat{fr}_{n,\bar{h}}^{LN2}$ are calculated in eqs 47, 48, and 49. The index set $\bar{H}_{(m)}$ includes the indices representing those system designs that, for at least one stage, the complementary design is part of $\{h_{(m-1)}^*(k), k \in K\}$. In other words, the system designs to be included in the partial design selection model (PDS) $_{(m)}$ are obtained by alternating the designs that are combinations of the stage designs in $\{h_{(m-1)}^*(k), k \in K\}$, which are represented by the dashed triangles in Figure 9b. Some parameters from the last step can be reused here.

$$\hat{C}_{\bar{h}}^U = \sum_{k \in K} \hat{C}_{k,h(k,\bar{h})}^U, \quad \forall \bar{h} \in \bar{H}_{(m)} \quad (46)$$

$$\widehat{fr}_{n,\bar{h}}^{LO2} = \sum_{k \in K} \widehat{FR}_{n,k,h(k,\bar{h}),\bar{h}(k,\bar{h})}^{LO2} = \sum_{k \in K} (\Theta_{n,k,h(k,\bar{h})}^{LO2} \cdot \Phi_{n,k,h(k,\bar{h})}^{LO2} + \Phi_{n,k,h(k,\bar{h})}^{LO2} \cdot \Theta_{n,k,h(k,\bar{h})}^{LO2}), \quad n \in N^{LO2}, \bar{h} \in \bar{H} \quad (47)$$

$$\widehat{fr}_{n,\bar{h}}^{LN2} = \sum_{k \in K} \widehat{FR}_{n,k,h(k,\bar{h}),\bar{h}(k,\bar{h})}^{LN2} = \sum_{k \in K} (\Theta_{n,k,h(k,\bar{h})}^{LN2} \cdot \Phi_{n,k,h(k,\bar{h})}^{LN2} + \Phi_{n,k,h(k,\bar{h})}^{LN2} \cdot \Theta_{n,k,h(k,\bar{h})}^{LN2}), \quad n \in N^{LN2}, \bar{h} \in \bar{H} \quad (48)$$

$$(PDS)_{(m)}: \min_{z_{\bar{h}}, x_n^{LO2}, x_n^{LN2}} C^U + C^T + PN$$

$$\text{s.t. } \sum_{\bar{h} \in \bar{H}_{(m)}} z_{\bar{h}} \geq 1, \quad \forall k \in K$$

$$C^U = \sum_{\bar{h} \in \bar{H}_{(m)}} z_{\bar{h}} \hat{C}_{\bar{h}}^U$$

$$\sum_{n \in N^{LO2}} x_n^{LO2} \geq 1$$

$$\sum_{n \in N^{LN2}} x_n^{LN2} \geq 1$$

$$C^T = \sum_{n \in N^{LO2}} x_n^{LO2} c_n^{LO2} + \sum_{n \in N^{LN2}} x_n^{LN2} c_n^{LN2}$$

$$f_{n,\bar{h}}^{LO2} \leq z_{\bar{h}} \widehat{fr}_{n,\bar{h}}^{LO2}, \quad \forall n \in N^{LO2}, \bar{h} \in \bigcup_{m' \leq m} \bar{H}_{(m')} \setminus \{h_{(m')}^*, m' \leq m-1\}$$

$$f_{n,\bar{h}}^{LN2} \leq z_{\bar{h}} \widehat{fr}_{n,\bar{h}}^{LN2}, \quad \forall n \in N^{LN2}, \bar{h} \in \bigcup_{m' \leq m} \bar{H}_{(m')} \setminus \{h_{(m')}^*, m' \leq m-1\}$$

$$PN = T(\text{penalty}^{LO2} \cdot \sum_{n \in N^{LO2}} \sum_{\bar{h} \in \bar{H}_{(m)}} f_{n,\bar{h}}^{LO2} + \text{penalty}^{LN2} \cdot \sum_{n \in N^{LN2}} \sum_{\bar{h} \in \bar{H}_{(m)}} f_{n,\bar{h}}^{LN2})$$

$$c^U, c^T, PN \geq 0$$

$$z_{\bar{h}} \in \{0,1\}, \quad \forall \bar{h} \in \bar{H}_k$$

$$x_n^{LO2} \in \{0,1\}, \quad \forall n \in N^{LO2}$$

$$x_n^{LN2} \in \{0,1\}, \quad \forall n \in N^{LN2} \quad (49)$$

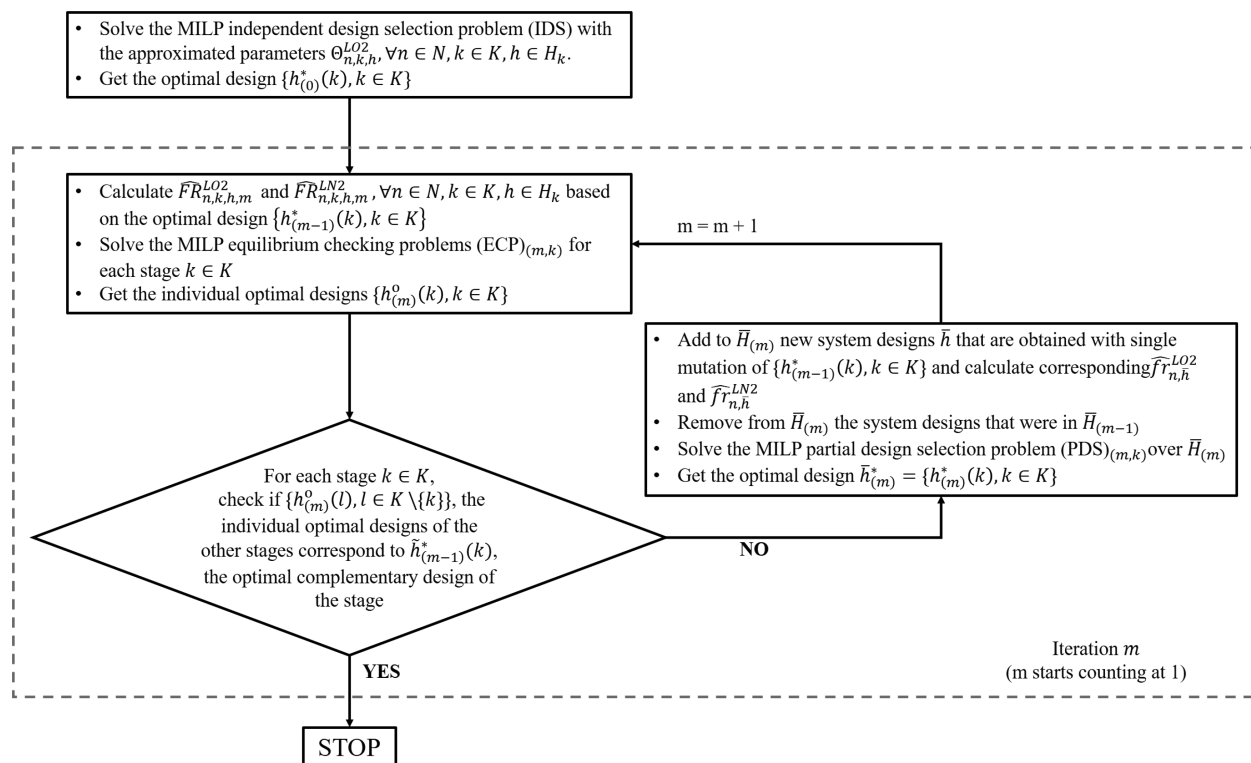


Figure 11. Flowchart of the iterative problem solving scheme.

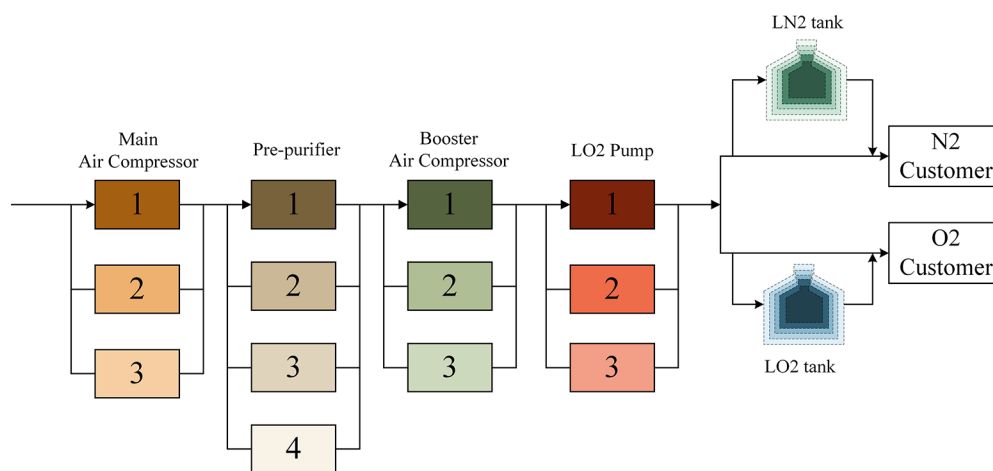


Figure 12. Design superstructure.

The optimal solution of $(PDS)_{(m)}$ indicates the potential design to select for each stage k : $\{h_{(m)}^*(k), k \in K\}$. After solving the model, let $m = m + 1$ and go back to step 2.

To reiterate, as shown in Figure 11, the algorithm iterates between two tasks, checking the equilibrium condition of $\{h_{(m)}^*(k), k \in K\}$, and finding a new design based on the unsuccessful previous solution and its superior deviation.

6. ADDITIONAL EXAMPLES

As was discussed in section 5.2, the globally optimal solution of the original formulation (RST) is a Nash-equilibrium solution but the reverse is not always true. However, we show in Appendix C that, given that the failure rates and repair rates of the units are of certain orders of magnitudes, the gap between an easily obtained initial design and the global optimum is

small and can be estimated, which makes it highly likely for the algorithm to arrive at the global optimum. In this section, we consider an example with the superstructure shown in Figure 12. It shares Table 2, Table 3, and Table 4 with the first example shown in section 4. Mean time between failure, mean time to repair, and capital costs are given in Table 1 in the Supporting Information, which are not real world values and are only meant for illustrative purposes. A series of reliability parameters sensitivity tests are performed, for which the game-theoretic approach gives the global optima according to the solution of the original MILP (RST) (model stats), but in much shorter time.

Figure 13 shows the optimal design. For both the main air compressor and the booster air compressor, units 2 and 3 are selected. The prepurifier and the LO2 pump exclude the last

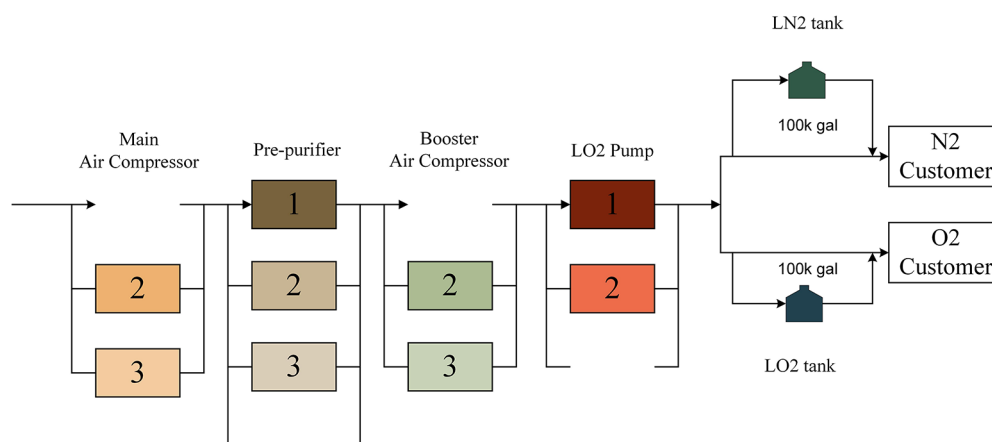


Figure 13. Optimal design for the parameters shown in Table 5.

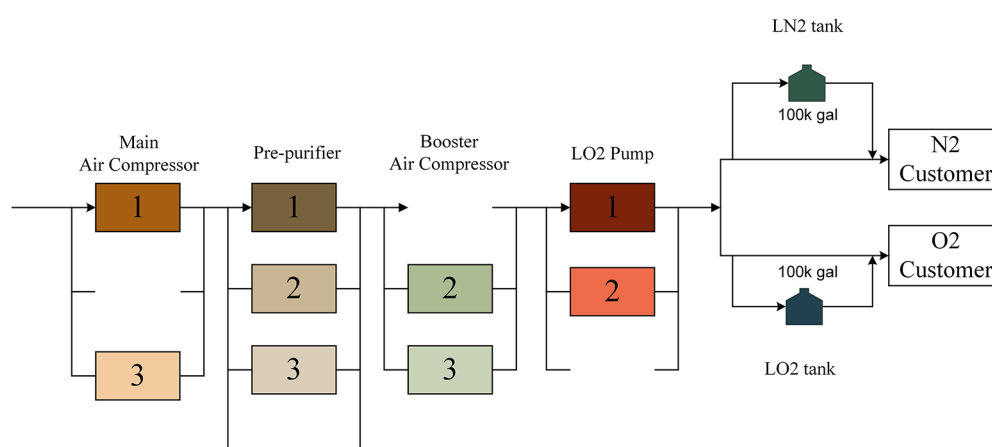


Figure 14. Optimal design when failure rates are doubled and repair rates are cut in half.

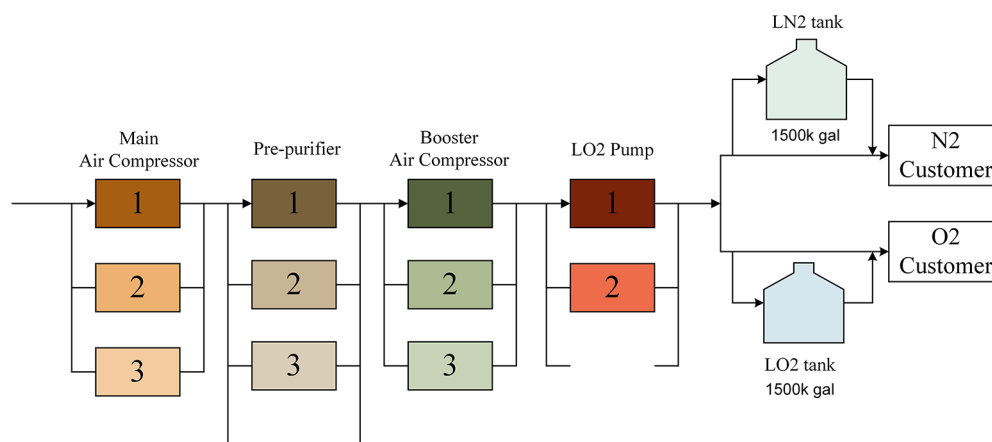


Figure 15. Optimal design when failure rates are multiplied by 5 and repair rates are divided by 5.

potential units. The tanks for both LN2 and LO2 are of size 100k gallon. The capital cost on the processing equipment is \$6,225k, while the capital cost on the storage tanks is \$105k. The expected frequency of LO2 outage is 0.0105 in the 10 year horizon, which incurs a \$21,089 penalty. The expected frequency of LN2 outage is 0.0121 in the 10 year horizon, which incurs a \$24,194 penalty.

First we perturb the parameters to make the units less reliable and harder to repair. When the failure rates are doubled and the repair rates are cut in half, the optimal design is shown in Figure 14,

where the only change compared to the original design (Figure 13) is that the main air compressor switched one of the redundancies for a more expensive and more reliable one. And because of the changes in the reliability parameters, the expected LO2 penalty becomes \$181,769, and the expected LN2 penalty becomes \$200,556.

If the failure rates are multiplied by 5 and repair rates are divided by 5 (Figure 15), the best decision is to have the largest storage tanks (1500k gallon) for both products, and to have all the potential units in the compressor stages. The capital cost on

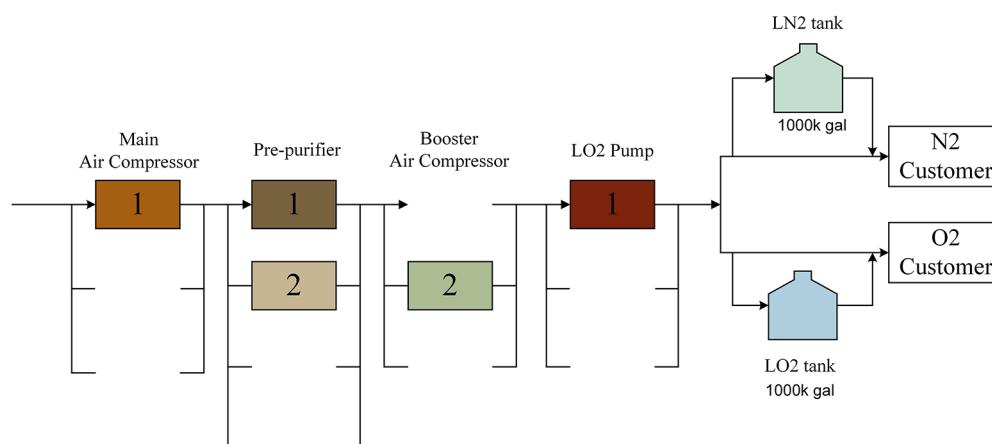


Figure 16. Optimal design when failure rates are cut in half and repair rates are doubled.

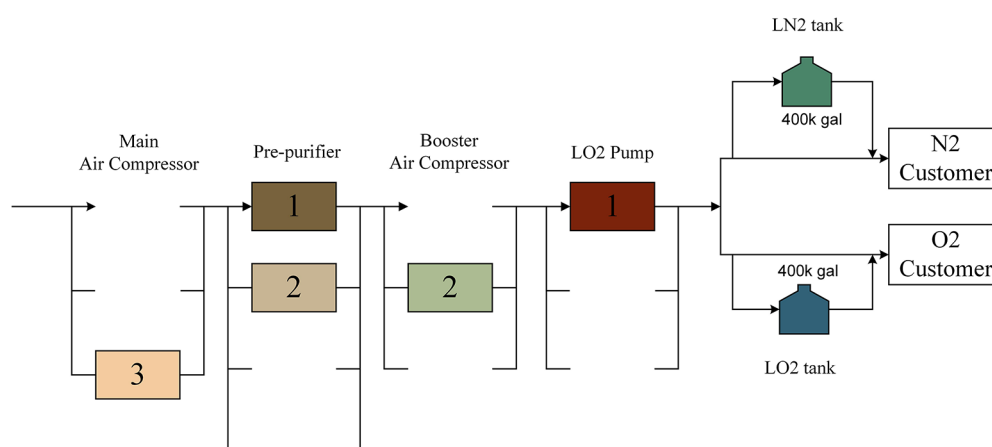


Figure 17. Optimal design when failure rates are divided by 5 and repair rates are multiplied by 5.

the processing equipment is now \$8,475k, while the capital cost on the storage tanks is \$1,815k. The expected frequency of LO2 outage is 0.254 in the 10 year horizon, which incurs \$508,643 penalty. The expected frequency of LN2 outage is 0.391 in the 10 year horizon, which incurs \$781,340 penalty.

Now we switch to the other direction where the failure rates are cut in half and the repair rates are doubled (Figure 16). There is a fairly large reduction in the capital investment needed for the redundancies comparing to the nominal case (\$3,415k compared to \$6,225k), but there has to be a raise on the storage side as a trade-off (1000k gallon for both LO2 and LN2 leading to \$1186k capital cost). The expected frequency of LO2 outage is 0.237 in the 10 year horizon, which incurs \$473,939 penalty. The expected frequency of LN2 outage is 0.295 in the 10 year horizon, which incurs \$589,553 penalty.

Finally, as shown in Figure 17 when the units are assumed to be significantly more reliable than the nominal case, the optimal design is to use the fewest number of redundancies and the smallest storage tanks. The investment on processing equipment is \$3,365k. 400k gallon tanks are chosen for both products, which leads to a \$452k cost. The expected frequency of LO2 outage is 0.0960 in the 10 year horizon, which incurs a penalty of \$192,029. The expected frequency of LN2 outage is 0.125 in the 10 year horizon, which incurs \$249,535 penalty.

It can be seen from the last two cases that the optimizer chooses to pay more penalties rather than higher investment cost, as the opportunistic decrease of outage penalties are not

Table 5. Numerical Results and Computational Statistics of the Sensitivity Test Cases

		least reliable	less reliable	more reliable	most reliable
objective value (K\$)		11579.9	6722.4	5855.6	4258.5
investment cost (K\$)		8475	6235	3415	3365
LN2 penalty (K\$)		508.6	181.8	666.8	192.1
LO2 penalty (K\$)		781.3	200.6	587.8	249.4
proposed algorithm	CPU	9.2	2.6	2.3	2.3
	no. iters.	3	1	1	1
original MILP (RST)	CPU for parameters	2702.2	2828.6	2712.3	2721.1
	CPU for models	1191.2	2198.2	604.6	189.1

enough to overcome capital commitment of installing more redundancies.

Table 5 shows the numerical results and the computational statistics from both methods for the above cases. It can be seen that the proposed game theoretic algorithm can solve the problem in much shorter time than the original MILP. Also, as the units become more reliable, the overall costs and penalties become lower.

7. CONCLUSION

In this paper, we have addressed the optimization problem of reliable design for chemical processes, in which design decisions

are made regarding redundant equipment selection and storage tanks sizing. The aim is to minimize the total cost, which consists of the penalty from product unavailability and the cost of increasing the reliability by installing parallel units and storage tanks. An MILP model (RST) based on Markov Chain assumption is proposed and applied to the motivating example of an air separation unit, where the pipeline availability is essential for ensuring the continuation of downstream customer productions.

While the Markov Chain framework closely represents the stochastic process of equipment failures and repairs, and the filling and consumption of liquid storage, the state space representation leads to the curse of dimensionality. To be able to tackle larger superstructures, we propose a game theoretic algorithm that decomposes and restructures the problem as a team game of the various processing stages, and arrives at a Nash Equilibrium among them. It is also shown that a good initial design that is close to the global optimum can be easily obtained for our problem, which guarantees the quality of the Nash Equilibrium solution. A number of examples have been shown to illustrate capability of the proposed algorithm to solve to global optimality in orders of magnitude shorter time than directly solving the original MILP model (RST).

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.iecr.9b04609>.

Three appendices of mathematical derivations, one appendix of algorithm illustration, and a long table as described in the text (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: grossmann@cmu.edu.

ORCID

Ignacio E. Grossmann: 0000-0002-7210-084X

Sivaraman Ramaswamy: 0000-0002-8495-8544

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank the National Science Foundation for support under Grant CBET-1705372, Praxair, Inc. (Now Linde plc), and Center for Advanced Process Decision-making at Carnegie Mellon University.

■ NOMENCLATURE

Indices

- k index of the processing stages
- h index of single stage designs
- \bar{h} index of the system designs
- \tilde{h} index of the minus-one-stage system designs
- s state index of single stages
- \bar{s} state index of the system
- \tilde{s} state index of the minus-one-stage systems
- n index of storage tank size options

Sets

- K index set of the processing stages
- H_k index set of the designs of stage k
- \bar{H} index set of system designs
- \tilde{H}_k index set of the complementary designs of stage k

- N^{LO2} index set of LO2 storage tank size options
- N^{LN2} index set of LN2 storage tank size options
- $S_{k,h}$ state space (set of all possible states) of design h in stage k
- $S_{k,h}^f$ set of all failure states of design h in stage k
- $\bar{S}_{\bar{h}}$ state space of the system design \bar{h}
- $\tilde{S}_{\tilde{h}}$ set of all failure states of system design \tilde{h}
- $\hat{S}_{k,\tilde{h}}$ state space of the complementary design \tilde{h} of stage k

Parameters

- $c_{\text{unit},k,h}$ investment cost of design h in stage k
- $W_{k,h}(s,s), \mathbf{W}_{k,h}$ transition rate (matrix) of design h in stage k
- $\pi_{k,h}(s), \boldsymbol{\pi}_{k,h}$ stationary probability distribution (vector) of design h in stage k
- $\sigma_{k,h}(s), \boldsymbol{\sigma}_{k,h}$ departing rate (vector) of design h in stage k
- $\bar{W}_{\bar{h}}(\bar{s},\bar{s}), \mathbf{W}_{\bar{h}}$ transition rate (matrix) of system design \bar{h}
- $\bar{\pi}_{\bar{h}}(\bar{s}), \bar{\boldsymbol{\pi}}_{\bar{h}}$ stationary probability distribution (vector) of system design \bar{h}
- $\bar{\sigma}_{\bar{h}}(\bar{s}), \bar{\boldsymbol{\sigma}}_{\bar{h}}$ departing rate (vector) of system design \bar{h}
- $f_{\bar{h}}(\bar{s})$ frequency of encountering state s of system design \bar{h}
- T planned service time of the ASU
- V_n^{LO2} reserve volume option n of liquid oxygen (tons)
- V_n^{LN2} reserve volume option n of liquid nitrogen (tons)
- $c_{\text{tank}_n}^{\text{LO2}}$ investment cost of liquid oxygen tank of reserve volume V_n^{LO2}
- $c_{\text{tank}_n}^{\text{LN2}}$ investment cost of liquid nitrogen tank of reserve volume V_n^{LN2}
- δ^{LO2} consumption rate of liquid oxygen (tons per day)
- δ^{LN2} consumption rate of liquid nitrogen (tons per day)
- penalty^{LO2} penalty of liquid oxygen supply interruption (K\$ per occurrence)
- penalty^{LN2} penalty of liquid nitrogen supply interruption (K\$ per occurrence)
- $\widehat{fr}_{n,\bar{h}}^{\text{LO2}}$ overconsumption frequency of liquid oxygen given that size n is selected for LO2 tank and \bar{h} is selected for system design
- $\widehat{fr}_{n,\bar{h}}^{\text{LN2}}$ overconsumption frequency of liquid nitrogen given that size n is selected for LN2 tank and \bar{h} is selected for system design
- $\widehat{FR}_{n,k,h(k,\bar{h})}^{\text{LO2}}$ overconsumption frequency of liquid oxygen that is due to the failure in stage k under system design \bar{h} , which corresponds to stage design h in stage k
- $\widehat{FR}_{n,k,h(k,\bar{h})}^{\text{LN2}}$ overconsumption frequency of liquid nitrogen that is due to the failure in stage k under system design \bar{h} , which corresponds to stage design h in stage k
- $\widehat{FR}_{n,k,h,\tilde{h}}^{\text{LO2}}$ overconsumption frequency of liquid oxygen that is due to the failure in stage k with design h and complementary design \tilde{h}
- $\widehat{FR}_{n,k,h(k,\tilde{h})}^{\text{LN2}}$ overconsumption frequency of liquid nitrogen that is due to the failure in stage k with design h and complementary design \tilde{h}

Variables

- $z_{k,h}$ binary variable that indicates the selection of design h for stage k
- x_n^{LO2} binary variable that indicates the selection of tank size option n for liquid oxygen
- x_n^{LN2} binary variable that indicates the selection of tank size option n for liquid nitrogen

$rt_{\bar{h}}(\bar{s})$	random variable of the residence time of state \bar{s} of system design \bar{h}
$V_{\bar{h}}^{\text{LO2dec}}(\bar{s})$	random variable of the volume of liquid oxygen decreasing during failure state \bar{s}
$V_{\bar{h}}^{\text{LN2dec}}(\bar{s})$	random variable of the volume of liquid nitrogen decreasing during failure state \bar{s}
$fr_{n,\bar{h}}^{\text{LO2}}$	overconsumption frequency of liquid oxygen given that size n is selected for LO2 tank and \bar{h} is selected for system design
$fr_{n,\bar{h}}^{\text{LN2}}$	overconsumption frequency of liquid nitrogen given that size n is selected for LN2 tank and \bar{h} is selected for system design
C^U	investment cost on processing units
C^T	investment cost on storage tanks
PN	expected penalty from supply interruptions

REFERENCES

- (1) Tan, J. S.; Kramer, M. A. A general framework for preventive maintenance optimization in chemical process operations. *Comput. Chem. Eng.* **1997**, *21*, 1451–1469.
- (2) Pistikopoulos, E. N.; Vassiliadis, C. G.; Arvela, J.; Papageorgiou, L. G. Interactions of maintenance and production planning for multipurpose process plants a system effectiveness approach. *Ind. Eng. Chem. Res.* **2001**, *40*, 3195–3207.
- (3) Cheung, K.-Y.; Hui, C.-W.; Sakamoto, H.; Hirata, K.; O'Young, L. Short-term site-wide maintenance scheduling. *Comput. Chem. Eng.* **2004**, *28*, 91–102.
- (4) Amaran, S.; Zhang, T.; Sahinidis, N. V.; Sharda, B.; Bury, S. J. Medium-term maintenance turnaround planning under uncertainty for integrated chemical sites. *Comput. Chem. Eng.* **2016**, *84*, 422–433.
- (5) Achkar, V. G.; Cafaro, V. G.; Méndez, C. A.; Cafaro, D. C. Discrete-Time MILP Formulation for the Optimal Scheduling of Maintenance Tasks on Oil and Gas Production Assets. *Ind. Eng. Chem. Res.* **2019**, *58*, 8231–8245.
- (6) Thomaidis, T.; Pistikopoulos, E. Integration of flexibility, reliability and maintenance in process synthesis and design. *Comput. Chem. Eng.* **1994**, *18*, S259–S263.
- (7) Thomaidis, T. V.; Pistikopoulos, E. Optimal design of flexible and reliable process systems. *IEEE Trans. Reliab.* **1995**, *44*, 243–250.
- (8) Gong, J.; You, F. Resilient design and operations of process systems: Nonlinear adaptive robust optimization model and algorithm for resilience analysis and enhancement. *Comput. Chem. Eng.* **2018**, *116*, 231–252.
- (9) Kuo, W.; Wan, R. *Computational intelligence in reliability engineering*; Springer, 2007; pp 1–36.
- (10) Aguilar, O.; Kim, J.-K.; Perry, S.; Smith, R. Availability and reliability considerations in the design and optimization of flexible utility systems. *Chem. Eng. Sci.* **2008**, *63*, 3569–3584.
- (11) Ye, Y.; Grossmann, I. E.; Pinto, J. M. Mixed-integer nonlinear programming models for optimal design of reliable chemical plants. *Comput. Chem. Eng.* **2018**, *116*, 3–16.
- (12) Ye, Y.; Grossmann, I. E.; Pinto, J. M.; Ramaswamy, S. Modeling for reliability optimization of system design and maintenance based on Markov chain theory. *Comput. Chem. Eng.* **2019**, *124*, 381–404.
- (13) Terrazas-Moreno, S.; Grossmann, I. E.; Wassick, J. M.; Bury, S. J. Optimal design of reliable integrated chemical production sites. *Comput. Chem. Eng.* **2010**, *34*, 1919–1936.
- (14) Lou, H. H.; Kulkarni, M.; Singh, A.; Huang, Y. L. A game theory based approach for emergy analysis of industrial ecosystem under uncertainty. *Clean Technol. Environ. Policy* **2004**, *6*, 156–161.
- (15) Castillo, L.; Dorao, C. A. Consensual decision-making model based on game theory for LNG processes. *Energy Convers. Manage.* **2012**, *64*, 387–396.
- (16) Zamarripa, M. A.; Aguirre, A. M.; Méndez, C. A.; Espuña, A. Mathematical programming and game theory optimization-based tool for supply chain planning in cooperative/competitive environments. *Chem. Eng. Res. Des.* **2013**, *91*, 1588–1600.
- (17) Radner, R. Team decision problems. *Ann. Math. Stat.* **1962**, *33*, 857–881.
- (18) Lewis, F. L.; Zhang, H.; Hengster-Movric, K.; Das, A. *Cooperative control of multi-agent systems: optimal and adaptive design approaches*; Springer Science & Business Media, 2013.
- (19) Marden, J. R.; Shamma, J. S. Game Theory and Control. *Annual Review of Control, Robotics, and Autonomous Systems* **2018**, *1*, 105–134.
- (20) Sericola, B. *Markov Chains: Theory and Applications*; John Wiley & Sons, 2013.
- (21) Trespalacios, F.; Grossmann, I. E. Review of mixed-integer nonlinear and generalized disjunctive programming methods. *Chem. Ing. Tech.* **2014**, *86*, 991–1012.
- (22) Young, P.; Zamir, S. *Handbook of game theory*; Elsevier, 2014.