Co-Scheduling of Energy and Production in Discrete Manufacturing Industrial Parks Considering Decision-Dependent Uncertainties

Yiyuan Pan, Zhaojian Wang

 \tilde{T}^h

Abstract—Modern discrete manufacturing industrial parks increasingly demand real-time energy and production coscheduling to minimize operational costs. The inherent complexity and diversity of production lines and equipment introduce significant uncertainties into the manufacturing process. Among these uncertainties, decision-dependent uncertainties (DDUs) present unique challenges, as the shape of the uncertainty set is influenced by decisions and cannot be determined prior to solving the model. However, existing research largely neglects the impact of DDUs in discrete manufacturing, and current algorithms addressing such uncertainties are computationally intensive, rendering them unsuitable for real-time industrial applications. To this end, this work develops an energy-production coscheduling model that explicitly incorporates decision-dependent uncertainties. Subsequently, multiple linearization techniques are proposed, each tailored to address the constraints associated with specific types of DDUs, enabling more efficient computation. Furthermore, a specialized algorithm, inspired by the columnand-constraint generation (C&CG) framework, is designed, along with a theoretical analysis of its convergence and computational complexity. The proposed method is validated through simulations on a real-world engine assembly line, demonstrating its capability to significantly reduce production costs while enhancing frequency regulation performance.

Index Terms—Decision-dependent uncertainty, independent uncertainty, discrete manufacturing.

NOMENCLATURE

A. Abbreviations

FK	Frequency regulation.
RTP	Real-time price.
DER	Distributed energy resource.
BESS	Battery energy storage system.

B. Variables

$I_{n,p}^h$	Operating state of equipment p in workshop n at moment h .
	at moment 70.
G_n^h	Output of workshop n moment h .
G_n^h C_n^h T^h	Input of workshop n moment h .
	Total time cost of equipment at moment h .
E^h	Total power consumption of equipment at mo-
	ment h .

This work was supported by the National Natural Science Foundation of China (62103265), and the Young Elite Scientist Sponsorship Program by the China Association for Science and Technology (No.YESS20220320). (Corresponding author: Zhaojian Wang)

Y. Pan and Z. Wang are with the Key Laboratory of System Control and Information Processing, Ministry of Education of China, Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China, (email:wangzhaojian@sjtu.edu.cn).

\tilde{E}^h	Actual power consumption at moment h .
$B_m^h \\ E_{EU}^h$	Inventory level in buffer m at moment h .
E_{EU}^{h}	Compensatory power extracted from BESS at
	moment h .
E_{LU}^h	Compensatory power extracted from DER at
	moment h .
$E^h_{SU} S^h$	Power stored from DER to BESS at moment h .
	BESS state of charge at moment h .
B_{ss}^h	Sales of by-products at moment h .

Actual time cost at moment h.

 \mathcal{K}

 \mathcal{R}_1

C. Parame	eters
$t_{n,p}$	Single use time cost of equipment p in workshop
	n.
$T_{(n,p)}$	Minimum single usage duration of equipment p
	in workshop n .
$e_{n,p}$	Single use power consumption of equipment p in
/1	workshop n .
$\alpha_{n,p}^h$	Product yield DDU of equipment p in workshop
70,12	n at moment h .
l_n	Workshop n 's location.
N_n	Daily maximum utilization limit for workshop n .
k_m	Single transport volume of buffer m .
Δt	Single transport time for buffer m .
M_s	Set of by-product output nodes.
$ar{S}$	Upper bound on BESS state of charge.
E_{fr}^{h}	Frequency regulation penalty DDU at moment h .
E_{fr}^h ζ^h	Product structure DDU at moment h .
s_m, s_s	Selling price of main and by-products.
E_{ex}^h	Expected frequency regulation requirements at
C.L	moment h .
J	Number of cell units in a battery.
V_{nom}	Nominal voltage of cell units in a battery.
β	Coefficients of battery degradation.
Δt^B	Battery degradation time interval.
$\underline{\alpha}_{n,p}$	Inherent lower bound of product yield of equip-

Distribution of E^h_{ex} derived from historical data. Distribution of E^h_{ex} affected by DDU. Line combinations and corresponding product structure at moment h derived from historical Probability of line combination i at hour h.

Set of specific production line combinations af-

ment p at moment h.

fecting yield.

I. Introduction

A. Backgroound

The discrete manufacturing sector has witnessed substantial growth in recent years, driven by the globalization of supply chains and the increasing demand for product customization [1]. Unlike process manufacturing, such as the chemical industry, discrete manufacturing—exemplified by engine assembly—typically involves a greater number of machines and more flexible, interchangeable process sequences. Simultaneously, the energy networks that power these machines, coupled with external distributed energy resources (DERs), create a complex and interconnected energy system. Thus, compared to process manufacturing, these characteristics result in a higher degree of freedom for a discrete manufacturing industrial park at the scheduling level, enabling more sophisticated decision-making and dynamic adjustments during production. [2]. Consequently, modern discrete manufacturing industrial parks have spurred increasing research interest in the coscheduling problem, which seeks to optimize the coordination of production and energy systems and reduce operational costs. However, the inherent uncertainties in the discrete manufacturing industry pose significant challenges to achieving efficient scheduling [3]–[5].

These uncertainties can be broadly categorized into two types: decision-independent uncertainties (DIUs) and decisiondependent uncertainties (DDUs). In the context of industrial parks, DIUs arise from the intrinsic characteristics of equipment or the production environment. Examples include variations in raw material quality and fluctuations in daily power generation from DER devices, such as solar panels. In contrast, DDUs are directly tied to production decisions. The main types of DDUs in industrial parks include the following: (1) The tendency to prioritize either main products or by-products depends on uncertain order statuses and future production conditions, which are, in turn, influenced by current production decisions. (2) The product yield is affected by both inherent equipment uncertainties and the selection of production lines. (3) The requirements for power frequency regulation, arising as industrial parks increasingly participate in grid services, are constrained by actual electricity usage decisions and linked to unpredictable demands from the power utility. The three types of DDUs will be elaborated upon in subsequent sections.

B. Related Works

Existing research on uncertainties in the scheduling processes of discrete manufacturing primarily addresses decision-independent uncertainties (DIUs). For instance, studies have focused on fluctuations in product and customer demand, proposing two-dimensional models to evaluate the system's ability to manage such uncertainties [6]. Other research investigates fluctuating daily production quantities and utilizes advanced algorithms, such as the non-dominated sorting adaptive differential evolution (NSJADE) algorithm, to derive robust order scheduling solutions [7]. Additionally, machine failures have been modeled using exponential distributions and handled through Markov chain-based methods [8]. The

uncertainty in machine yield, treated as a fuzzy variable and characterized by a probability-of-failure-working-time equation, has also been analyzed in depth [9]. Another line of research examines performance degradation, such as equipment failures, employing a combined approach of Adaboost, DNN, and LSTM for simultaneous updates [10]. While these studies provide valuable insights into DIUs, some uncertainties, such as equipment defect rates [9], [10], which are inherently linked to both the intrinsic characteristics of the equipment and the interactions with upstream production processes, should instead be modeled as decision-dependent uncertainties (DDUs) to more accurately capture their impact on scheduling outcomes.

However, modeling DDUs introduces additional challenges due to the coupling between decisions and uncertainties, which significantly complicates the optimization problem. To address these challenges, researchers have developed advanced algorithms for solving optimization models that incorporate DDUs. Existing methodologies often rely on iterative decomposition algorithms or scenario tree-based partitioning techniques. Multi-stage models with DDUs have been addressed using recursive optimization methods combined with numerical enumeration techniques [11]. Polyhedral representations of DDUs have been explored using iterative algorithms based on Benders decomposition, where optimality and feasibility cuts mitigate the coupling between uncertainty and decisionmaking [12]. Other approaches transform optimization problems with DDUs into dualized forms and apply adversarial cutting-plane algorithms for solution derivation [13]. Scenario tree frameworks have been utilized in conjunction with model predictive control to decompose DDU-based optimization problems into subproblems, solving them under different learning scenarios [14]. In addition, adaptive reliability improvement unit commitment (ARIUC) algorithms have been proposed to efficiently decompose DDU-related models into tractable subproblems [15]. While these methodologies significantly advance the understanding of DDUs, both iterative decomposition and scenario tree-based partitioning techniques suffer from high computational complexity. This limitation poses challenges for their application in real-time and rapid scheduling scenarios, which are increasingly critical in modern industrial contexts.

C. Methodology

To summarize, existing studies on the co-scheduling of industrial production and energy under uncertainty exhibit two key challenges in modeling and solving: (1) *Modeling*: To the best of our knowledge, existing research primarily addresses DIUs while overlooking DDUs. However, certain DIUs are intrinsically tied to decision-making processes and should be modeled as DDUs for more comprehensive representations. 2) *Solving*: Traditional methodologies for handling DDU-related problems focus exclusively on solvability. In modern industrial contexts, there is a critical demand for algorithms with low computational complexity to enable rapid and real-time control, thereby achieving optimal production outcomes.

To address these issues, this paper proposes a two-stage robust optimization model that accounts for a complete set of three aforementioned types of DDUs in discrete manufacturing, accompanied by a novel algorithm with performance guarantees. First, a robust optimization model incorporating multiple DDUs is constructed. The constraints associated with the three types of DDUs are then linearized using methods based on ambiguity set theories, the imprecise Dirichlet model, and Cantelli's inequality. By incorporating these techniques alongside additional simplification methods such as dual transformation, the problem is reduced to a mixedinteger linear programming (MILP) problem. Building on this reformulation, we proposed the decision-dependent C&CG (DDCCG) algorithm based on the traditional C&CG algorithm and provided theoretical performance analyses to establish its convergence and optimality. Finally, a case study involving a real-world engine assembly line modeled using Petri nets is presented. The results demonstrate significant cost reductions, enhanced anti-interference capabilities of the production line, and improved responsiveness to frequency regulation and peak shaving requirements. These findings validate the efficacy of the proposed model and algorithm.

The key contributions of this paper are summarized as follows:

- This paper presents a two-stage robust optimization model for the discrete manufacturing industry that considers multiple DDUs. In particular, the model incorporates product yield DDU, frequency regulation penalty DDU, and product structure DDU into the energy and production co-scheduling process. Compared to existing models, it more accurately captures the interplay between decisions and uncertainties in real-world industrial park operations.
- We presented three paradigms for reducing different DDU-related constraints to a linear form. In particular, the three types of DDUs in discrete manufacturing are mathematically represented as ambiguity sets, univariate distributions, and multivariate distributions, which encompass the majority of real-world DDUs. Therefore, the proposed methods in this paper are broadly applicable across diverse engineering scenarios.
- We proposed the DDCCG algorithm for the decisiondependent two-stage robust optimization model and proved its convergence and optimality theoretically. The original C&CG algorithm has a fast convergence speed but cannot adapt to the decision-dependent model. Our DDCCG algorithm ensures fast convergence and optimality while meeting the demands of modern industrial systems.

The remainder of this paper is structured as follows. Section II introduces the foundational optimization problem. Section III details the modeling and linearization techniques for DDUs. Section IV elaborates on problem simplification and algorithm design. Section V provides a comprehensive case study analysis. The final section summarizes the work.

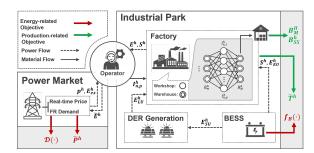


Fig. 1: Discrete manufacturing industrial park architecture.

II. PROBLEM FORMULATION

A. Problem Overview

A modern discrete manufacturing industrial park typically comprises a production system and an energy system, see Fig 1. Without loss of generality, the production system consists of production workshops and warehousing facilities [16], which can be represented as a directed graph G := (V, E), where $V = \{1, ..., N; N + 1, ..., N + M\}$ denotes the set of nodes. N represents the number of production workshops, and Mrepresents the number of warehousing facilities. The directed edge $E \subseteq V \times V$ captures the flow of materials and products between nodes. The production workshops provide various services, while the warehousing facilities facilitate the transfer of semi-finished products within the system. For simplicity, the energy network includes a DER generation unit and a battery energy storage system (BESS), both of which are connected to the production network. Integrated with the production network. These components allow the operator to access and manage energy resources to power the production equipment effectively. The scheduling horizon is defined as $T = \{1, ..., H\}$, over which production and energy decisions are jointly optimized.

In our paper, the scheduling process is modeled as a twostage robust optimization model and the operator of the industrial park will make decisions based on data collected from the power grid (e.g., electricity prices, frequency regulation requirements) and production information from the factory. The objective is to maximize production profitability while ensuring the stability of the energy system.

B. Objective Function

As the entity responsible for maintaining the industrial park, our primary objective is to develop strategies that minimize the total economic cost. The objective function is formulated as follows:

$$\min_{\substack{I_{n,p}^{h}, E_{EU}^{h}, E_{LU}^{h}} \sum_{h=0}^{H} \tilde{T}^{h} \cdot r + f_{B}(S^{h}) \\
+ \sum_{h=0}^{H} \tilde{P}^{h} + \sum_{h=0}^{H} E_{fr}^{h} \\
- B_{M}^{H} \cdot s_{m} - \zeta^{h} \sum_{h=0}^{H} B_{ss}^{h} \cdot s_{s} \quad (1)$$

In (1), The first term represents the equipment usage loss (intuitively, time of use and the maintenance costs are proportional.), the second term captures the battery degradation

penalty [17], [18], the third term corresponds to the power purchase cost from the grid, the fourth term is the frequency regulation penalty term, reflecting deviations from specified grid requirements [19], and the fifth and sixth terms represent the revenue from the sale of main products and by-products, respectively. The specific formulation of the battery degradation function is provided in (2a) and (2b), following the models detailed in [17], [18].

$$f_B(S^h) = J\lambda_{cell}V_{nom}h_{cell}(S^h)\Delta t^B, \forall t$$
 (2a)

$$h_{cell}(S^h) = \beta S^h, \forall t$$
 (2b)

Remark 1. (DDU variable ζ^h) The factory's production preference for main and by-products is modeled as a DDU, characterized by the weight factor ζ^h . This variable reflects the factory's production tendencies, which are influenced by the inherent uncertainty in future order expectations and production dynamics. However, influenced by actual production conditions, these tendencies exhibit decision-dependent behavior, making ζ^h a critical DDU in the model.

C. Production Model and Constraint

Without loss of generality, the production-related variables for workshop n at time h can be described using a vector representation, $[\mathrm{ID}_n, I_{n,p}^h, G_n^h, C_n^h, t_{n,p}, e_{n,p}, \alpha_{n,p}, \tilde{l}_n]$, where ID_n denotes the workshop ID, $I_{n,p}^h$ is a 0-1 variable representing the state of the p-th equipment or service in workshop n at time $h,\,G_n^h$ and G_n^h represent the manufacturing output and input of workshop n at time h, $t_{n,p}$ and $e_{n,p}$ are the manufacturing time and energy consumption of workshop n at time h, $\alpha_{n,p}$ denotes the yield rate of the p-th equipment or service in workshop n, and k_n indicates the geographic location of workshop n. These variables collectively capture the essential aspects of the production system and must satisfy a set of predefined constraints to ensure feasibility and operational efficiency.

$$\sum\nolimits_{p=1}^{P} I_{n,p}^{h} \le 1 \tag{3a}$$

$$\sum_{p=1}^{P} I_{n,p}^{h} \le 1$$

$$\sum_{h=1}^{H} \sum_{p=1}^{P} I_{n,p}^{h} \le N_{n}$$
(3a)
(3b)

$$T^{h} = \sum_{n=1}^{N} \sum_{p=1}^{P} I_{n,p}^{h} \cdot t_{n,p}$$
 (3c)

$$E^{h} = \sum\nolimits_{n = 1}^{N} {\sum\nolimits_{p = 1}^{P} {I_{n,p}^{h} \cdot e_{n,p} } }$$
 (3d)

$$\tilde{T}^h = T^h + \frac{B_m^h - B_M^{h-1}}{k_m} \cdot \Delta t \tag{3e}$$

$$G_n^h = \sum_{p=0}^P I_{n,p}^h \cdot g_{n,p} \cdot \alpha_{n,p}^h \tag{3f}$$

$$C_n^h = \sum_{p=0}^P I_{n,p}^h \cdot c_{n,p} \tag{3g}$$

Constraint (3a) ensures the uniqueness of the service assigned to workshop n at moment h, preventing overlap or redundancy. Constraint (3b) imposes a daily maximum usage limit for each piece of equipment to maintain operational sustainability. Constraints (3c) and (3d) define the production time and energy costs at time h. Constraint (3e) specifies that the production time cost comprises both transportation time and equipment operation time, reflecting a comprehensive

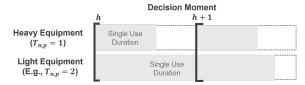


Fig. 2: Minimum usage time constraint for heavy equipment to ensure operational continuity and cost efficiency.

cost structure. Finally, constraints (3f) and (3g) quantify the production output and input for workshop n at time h, ensuring consistency and balance in the production system.

Remark 2. (DDU Variable $\alpha_{n,p}^h$) The production yield $\alpha_{n,p}^h$, influenced by the intrinsic uncertainties associated with the equipment, is modeled as a DDU. Additionally, the yield of specific equipment is affected by the interdependence of upstream processes or combinations of production lines. For instance, in engine assembly, the success rate of the crankshaft mounting process is significantly improved when the semifinished products have undergone high-quality crankshaft grinding. This decision-dependent nature highlights the impact of prior process decisions on subsequent production outcomes.

Existing work adopts a time-rolling-based approach, which is incompatible with actual production processes in discrete manufacturing dispatch. Factories are inherently discrete-time systems, and the time-rolling base assumes that the production states of neighboring moments, $I_{n,p}^h$ and $I_{n,p}^{h+1}$, are independent and operators can ascertain the production status of all equipment at each moment. However, this approach neglects the duration of equipment utilization, which varies across different pieces of equipment, and fails to capture the operational constraints of large machines with longer minimum usage times. Such machines are costly to start up, resulting in interdependencies between production states at consecutive time steps. As illustrated in Figure 2, when heavy equipment is activated at time t, it will remain operational at time t+1for economic reasons. To address this limitation, we introduce the following constraint that enforces a minimum usage time $T_{n,p}$ for equipment p, ensuring operational continuity and cost efficiency.

$$\sum_{k \in P(n)} I_{n,k}^h \cdot (I_{n,k}^{h+t} - I_{n,p}), \forall t = 0, 1, ..., T_{n,p}$$
 (4)

Having analyzed the constraints related to the production line, we now turn to the warehousing and logistics aspects. We assume that buffers are positioned at the back end of each workshop to facilitate the storage and distribution of semifinished products. These buffers impose additional constraints on the production system, which we derive as follows.

$$B_m^{h+1} = B_m^h + \Delta W_m^{h+1}, m \notin M_s$$
 (5a)

$$B_m^{h+1} = B_m^h - B_{ss}^h + \Delta W_m^{h+1}, m \in M_s$$
 (5b)

$$\Delta W_m^{h+1} = \sum_{i \in P(m)} G_i^h - \sum_{i \in S(m)} C_i^h$$
 (5c)

$$B_m^h \ge 0 \tag{5d}$$

$$B_m^0 = \begin{cases} C, & \text{if } m = 0, \\ 0, & \text{if } m \neq 0. \end{cases}$$
 (5e)

Equations (5a) and (5b) define the time recurrence relationships for the buffers, where constraint (5b) specifically addresses special buffers that are capable of directly outputting by-products. Equation (5c) models the changes in inventory levels, where W_m^{h+1} represents the difference between the production of upstream workshops and the consumption by downstream workshops for buffer m. The functions $P(\cdot)$ and $S(\cdot)$ denote the sets of upstream and downstream workshops concerning the current buffer, respectively. Finally, equations (5d)–(5e) establish the boundary conditions governing the storage states, ensuring feasibility and operational consistency within the system.

D. Energy Model and Constraints

The integration of DER devices and BESSs into modern industrial parks aims to reduce electricity procurement costs. For simplicity, this paper focuses on solar panels as the representative DER device [20]. The inclusion of BESSs and DER devices introduces the following constraints for the industrial park's operation.

$$\tilde{E}^h = E^h - E^h_{EU} - E^h_{LU} \tag{6a}$$

$$E_{DER}^{h} = E_{SU}^{h} + E_{LU}^{h}, \ \underline{E}_{DER}^{h} \le E_{DER}^{h} \le \bar{E}_{DER}^{h}$$
 (6b)
$$\tilde{P}^{h} = \tilde{E}^{h} \cdot p^{h}$$
 (6c)

$$\tilde{P}^h = \tilde{E}^h \cdot p^h \tag{6c}$$

$$S^{h+1} = S^h - \lambda \cdot E_{EU}^h + \xi \cdot E_{SU}^h$$

$$0 \le S^h \le \bar{S}$$
(6c)
$$(6d)$$

$$(6d)$$

$$0 < S^h < \bar{S} \tag{6e}$$

$$k < |S^{h+1} - S^h| < \bar{k} \tag{6f}$$

$$E_{fr}^h \ge \mathfrak{D}(\tilde{E}^h - E_{ex}^h) \tag{6g}$$

Equation (6a) specifies the nodal power injection constraint, ensuring a balance between supply and demand at each node. Equation (6b) represents the allocation of power generated by the DER, which can either be stored in the BESS or consumed directly, where E_{DEB} , the DER generation, is a DIU. Equation (6c) calculates the total electricity procurement cost, incorporating real-time price (RTP) based on peak and valley tariffs. Equations (6d)-(6f) define the constraints associated with the BESS. Specifically, (6d) models the state transfer dynamics of the BESS, while (6e) and (6f) impose limits on the energy storage capacity and enforce adjacent time interval creep constraints. The parameters k and k denote the lower and upper bounds of the energy storage system, respectively. Finally, equation (6g) introduces the frequency regulation penalty term, where $\mathfrak{D}(\cdot)$ represents a general distance measure function to quantify deviations.

Remark 3. (DDU variable E_{fr}^h) The penalty term for frequency regulation, E_{fr}^h , is influenced by the expected power consumption E_{ex}^{h} from the power utility, which remains unknown to the factory. Additionally, it depends on the factory's actual power consumption data and other uncertain variables. Due to the interdependence between decisions and these uncertainties, E_{fr}^h is classified as a DDU.

III. MODELING OF DDUS

This section examines the three most prominent forms of DDUs encountered in discrete manufacturing industrial parks. The mathematical representations of these DDU variables are provided, followed by their simplification using different linearization techniques.

A. Classification of DDUs

Depending on DDUs' characteristics and the decisionmaker's state of knowledge, DDUs can be categorized into two distinct forms. If the decision-maker can estimate the relationship between DDUs and their corresponding decision variables based on historical data or physical laws, a probability distribution is employed to characterize the uncertainty. Conversely, if no prior knowledge is available, DDUs are represented using ambiguity sets. For probability-based DDUs, further categorization is possible into univariate and multivariate probability distributions, reflecting the dimensionality of the uncertainties.

In the context of discrete manufacturing production, the aforementioned types of DDUs can be classified as follows.

- Product yield DDU: The product yield DDU is influenced by the decision variables associated with different production line combinations. However, the decisionmaker lacks precise knowledge of the uncertainty associated with the equipment in each line. Consequently, this DDU is best represented using rough boundaries and falls under the category of ambiguity set DDUs.
- Frequency regulation penalty DDU: The frequency regulation penalty DDU is influenced by the frequency regulation demand from the power utility, which is unknown to the factory. However, decision-makers can estimate the expected values of this DDU based on historical data, allowing for a univariate probabilistic characterization.
- Product Structure DDU: Decision-makers can estimate tendencies for different products based on historical order information. However, the historical data reflects various factors, including production line combinations and specific equipment choices. As a result, this DDU requires a multivariate probability distribution model for representation. Without loss of generality, this paper adopts the imprecise Dirichlet model to characterize this DDU.

B. Definition and Simplification of DDUs

1) **Product Yield DDU**: As previously discussed, the combination of production lines significantly impacts the yield of a given product. For instance, if a specific production line combination incorporates additional repair processes for semiproducts, the yield is likely to improve. To account for these effects, the yield rate $\alpha_{n,p}^h$ is redefined as follows.

$$\alpha_{n,p}^h \ge \underline{\alpha}_{n,p} - d_{n,p}(I_{n,p}^h), p \in \mathcal{P}$$
 (7a)

$$d_{n,p}^{h} = \sum_{k \in \mathcal{K}} I^{k,h} \Delta \alpha_k \tag{7b}$$

$$I^{k,h} = \prod_{k, \in k} I^{h}_{n,k_1} I^{h}_{n,k_2} \dots I^{h}_{n,k_N}$$
 (7c)

In (7a), the correction parameter $d_{n,p}$ represents the influence of production line combinations, as detailed in equation (7b). The parameter $\underline{\alpha}_{n,p}$ denotes the lower bound of the yield inherent in the equipment. The set \mathcal{P} comprises the equipment subject to correction, while the set K includes the production line combinations that influence the finished product rate. These re-definitions ensure a more accurate characterization of product yield under decision-dependent relationships.

2) Frequency Regulation Penalty DDU: For the expected power consumption E_{ex}^h in the frequency regulation term of the objective function, we first define the general frequency regulation distance measure as follows. The function $\mathfrak D$ denotes a generic distance measure, and in this work, we adopt the L-1 norm \mathcal{D} . The frequency regulation penalty term is then expressed as:

$$E_{fr}^h \ge \mathcal{D}(\tilde{E}^h - E_{ex}^h) \tag{8}$$

The historical expectation data E_{ex}^h is modeled using a Gaussian distribution, as described in equations (9a)-(9c). Suppose the dataset for time h contains T samples, the mean μ_h and variance σ_h^2 are calculated as:

$$\mu_h = \frac{1}{T} \sum_{n=1}^{T} E_{ex,n}^h$$
 (9a)

$$\sigma_h^2 = \frac{1}{T} \sum_{n=1}^{T} (E_{exp,n}^h - \mu_h)^2$$
 (9b)

$$\sigma_h^2 = \frac{1}{T} \sum_{n=1}^T (E_{exp,n}^h - \mu_h)^2$$

$$\mathcal{R}_1 = \left\{ \mathbb{P}_r \in \mathcal{P}(\mathbb{R}) : \\ \mathbb{E}_{\mathbb{P}_r}[(E_{fr}^h - \mu^h)^2] = \sigma_h^2 \right\}$$
(9b)

In this formulation, μ_h and σ_h represent the mean and standard deviation of the historical data. To account for fluctuations in power consumption, we extend the Gaussian distribution to include adjustments for these variations. The modified mean $\mu_h(\tilde{E}^h)$ and the revised probability distribution constraints are given as:

$$\mu_h(\tilde{E}^h) = \mu_h - (K \cdot \tilde{E}^h + B)$$

$$(10a)$$

$$\mathcal{R}_2 = \left\{ \mathbb{P}_r \in \mathcal{P}(\mathbb{R}) : \mathbb{E}_{\mathbb{P}_r}[E_{fr}^h] - \mu_h(\tilde{E}^h))^2 \cdot \sigma_h^{-1} \leq \gamma_1 \right\}$$

$$\mathbb{E}_{\mathbb{P}_r}[(E_{fr}^h - \mu^h)^2] \leq \gamma_2 \cdot \sigma_h$$

$$(10b)$$

The constraint \mathcal{R}_2 is further simplified using linearization techniques. First, we rewrite the constraint as:

$$\inf_{\mathbb{P}_r \in \mathcal{R}_2} \mathbb{P}_r(E_{fr}^h \le C) \ge 1 - \epsilon, \tag{11}$$

where C is a constant, and ϵ denotes a predefined confidence level. Following the approach in [21], auxiliary variables $\tilde{s}=E^h_{fr}-\mu_h(\tilde{E}^h)$ and $b=C-\mu_h(\tilde{E}^h)$ are introduced, along with the sets:

The probability distribution inscription has been revised as (10). After obtaining \mathcal{R}_2 , a linearization step is needed for simplification. We first rewrite the constraint in the form (12).

$$\inf_{\mathbb{P}_r \in \mathcal{R}_2} \mathbb{P}_r(E_{fr}^h \le C) \ge 1 - \epsilon \tag{12}$$

$$S = \{(\mu_1, \sigma_1) : |\mu_1| \le \sqrt{\gamma_1 \sigma_h}, \, \mu_1^2 + \sigma_1^2 \le \gamma_2 \sigma_h\}, \quad (13a)$$

$$\mathcal{R}_{2} = \left\{ \begin{aligned} & |\mathbb{E}_{\mathbb{P}_{r}}[\tilde{s}]| \leq \sqrt{\gamma_{1}} \sqrt{\sigma_{h}}, \\ & \mathbb{E}_{\mathbb{P}_{r}}[\tilde{s}^{2}] \leq \gamma_{2} \cdot \sigma_{h}, \quad \mathbb{P}_{r} \in \mathcal{P}(\mathbb{R}). \end{aligned} \right\}$$
(13b)

The DDU constraint is thus decomposed into a two-layer optimization problem as (14):

$$\inf_{\mathbb{P}_r \in \mathcal{R}_2} \mathbb{P}_r(E_{fr}^h \le C) = \inf_{\mathbb{P}_r \in \mathcal{D}_{\tilde{s}}} \mathbb{P}_r\{\tilde{s} \le b\} = \inf_{(\mu_1, \sigma_1) \in S} \inf_{\mathbb{P}_r \in \mathcal{R}_1} \mathbb{P}_r\{\tilde{s} \le b\} \quad (14)$$

For the outer layer, the worst-case condition is identified, while the inner layer determines the worst-case distribution constrained by the given mean and variance. Using Cantelli's inequality [21], the constraint is rewritten as:

$$\inf_{\mathbb{P}_r \in \mathcal{R}_2} \mathbb{P}_r(E_{fr}^h \le C) \le \inf_{(\mu_1, \sigma_1) \in S} \frac{(b - \mu_{0,h})^2}{\sigma_{0,h}^2 + (b - \mu_{0,h})^2}$$
 (15)

Given that \mathbb{P}_r and the parameters in S are known from historical data as defined, the constraint is further linearized. Representing \mathbb{P}_r 's mean and variance as $\mu_{\mathbb{P}_r}$ and $\sigma_{\mathbb{P}_r}$, respectively, we rewrite the constraint as:

$$E_{fr}^{h} - \mu_{\mathbb{P}_r} - \sigma_{\mathbb{P}_r} \Phi^{-1} \left(\frac{(b - \mu_{0,h})^2}{\sigma_{0,h}^2 + (b - \mu_{0,h})^2} \right) \le 0$$
 (16)

where Φ is the cumulative distribution function (CDF) of \mathbb{P}_r . Substituting the decision variables, the optimization constraints are finalized as:

$$-\Delta E^h + E^h_{ex} \le \tilde{E}^h \le \Delta E^h + E^h_{ex} \tag{17a}$$

$$\Delta E^{h} = \mathcal{D}^{-1} \left[\mu_{\mathbb{P}_r} - \sigma_{\mathbb{P}_r} \Phi^{-1} \left(\frac{(b - \mu_{0,h})^2}{\sigma_{0,h}^2 + (b - \mu_{0,h})^2} \right) \right]$$
(17b)

Here, \mathcal{D}^{-1} denotes the inverse of the distance measure function. Since \mathcal{D} and the values in \mathcal{D}^{-1} are derived from historical data, constraint (17a) is linear and solvable.

3) **Products Structure DDU**: The operator's tendency to prioritize different products, influenced by their varying profit margins, is modeled as ζ^h , a multivariate DDU. Since this tendency is partially known from historical sales data, we represent it using a probability distribution as follows:

$$D^{h} = \{LC_{1}^{h}, LC_{2}^{h}, ..., LC_{K}^{h}\}$$
 (18a)

$$W^h = \{w_1^h, w_2^h, ..., w_K^h\}$$
 (18b)

$$\mathbb{P}_r(w_i^h) = \mathbb{P}_r(LC_i^h) = \theta_i^h, i = 1, 2, ..., K$$
 (18c)

$$f(\theta^h) = \Gamma(s) [\prod\nolimits_{i=1}^n \Gamma(s \cdot r_i)]^{-1} \prod\nolimits_{i=1}^n \theta_i^{s \cdot r_i - 1} \qquad (18\mathrm{d})$$

$$\sum_{i=1}^{n} r_i = 1, \quad \forall r_i \in [0, 1]$$
 (18e)

Here, the set D^h contains K production line combinations LC_i^h derived from historical data, and W^h represents the byproduct to main product sales ratio w_i^h for each combination, computed as the average from historical observations. A higher w_i^h indicates a stronger tendency toward by-products. The Imprecise Dirichlet Model (IDM) [22] is employed to inscribe the probability distribution, where Γ , r_i^h , θ_i^h , and s represent the Gamma function, the prior weight factor for the i-th state, the probability of the *i*-th state, and the equivalent sample size (assumed to be 1), respectively.

Based on Bayesian theory, the posterior distribution of θ^h follows a Dirichlet distribution, represented as (19).

$$N^h = n_1^h + n_2^h + \dots + n_K^h \tag{19a}$$

$$n_i = n_{i,\text{HD}}^h + n_{i,\text{RT}}^h \tag{19b}$$

$$f(\theta^{h}|N) = \Gamma(s+N^{h}) \left[\prod_{i=1}^{n} \Gamma(s \cdot r_{i} + n_{i}) \right]^{-1} \prod_{i=1}^{n} \theta_{i}^{s \cdot r_{i} + n_{i} - 1}$$
(19c)

In this formulation, N^h is the total number of observations, combining real-time production data at moment h with historical data. The parameter n^h_i includes contributions from the historical dataset $(n^h_{i,\mathrm{HD}})$ and real-time data $(n^h_{i,\mathrm{RT}})$. Using this posterior density function, we derive the uncertainty interval of the probability as follows:

$$\zeta^h = \sum_{i=1}^K \theta_i^h \tag{20a}$$

$$\theta_i^h = [\underline{E}(\theta_i^h), \bar{E}(\theta_i^h)] = \left[\frac{m_i}{s+K}, \frac{m_i+s}{s+K}\right]$$
 (20b)

To ensure robustness, we calibrate the uncertainty set using a confidence level γ , defining the upper and lower bounds of θ_i^h as (21).

$$\begin{cases} \frac{\underline{\theta}^h = 0, \bar{\theta}^h = G^{-1}(\frac{1+\gamma}{2}), n_i^h = 0\\ \underline{\theta}^h = H^{-1}(\frac{1-\gamma}{2}), \bar{\theta}^h = G^{-1}(\frac{1+\gamma}{2}), 0 \le n_i^h \le N^h\\ \underline{\theta}^h = H^{-1}(\frac{1-\gamma}{2}), \bar{\theta}^h = 1, n_i^h = N^h \end{cases} \tag{21}$$

In (21), γ , H, and G represent the confidence coefficient, the cumulative distribution function (CDF) of $B(m_i^h,s+N^h-m_i^h)$, and the CDF of $B(m_i^h+s,N^h-m_i^h)$, respectively. The interval of θ^h is defined by $\underline{\theta}^h$ and $\bar{\theta}^h$. The resulting constraints on ζ^h are expressed as:

$$\zeta^h = \sum_{i}^{K} \theta_i^h(I_{n,p}^h) \tag{22a}$$

$$\underline{\theta}^h \le \theta_i^h(I_{n,n}^h) \le \bar{\theta}^h \tag{22b}$$

It is worth noting that the bounds $\underline{\theta}^h$ and $\overline{\theta}^h$ in (22b) are dynamically updated during each iteration based on n_i^h following (21). However, the constraints remain linear within the same loop. More extensive historical data leads to narrower probability intervals, enhancing the optimization's fidelity to real-world conditions [23].

IV. PROBLEM TRANSFORMATION & ALGORITHM DESIGN

In this section, we derive the final formulation of the optimization problem and decompose it into the corresponding master problem (MP) and sub-problem (SP). Subsequently, we propose a novel algorithm specifically designed to solve this model efficiently. Theoretical analyses of the proposed algorithm are presented, focusing on its convergence properties and computational complexity.

A. Problem Transformation

The final optimization problem is expressed as follows.

$$\min_{I_{n,p}^h} \sum_{h=0}^H \tilde{T}^h + f_B(S^h) + \sum_{h=0}^H \tilde{P}^h$$

$$+ \max_{\text{DDU,DIU}} \min_{E_{EU}^h, E_{LU}^h} B_M^H s_m - \zeta^h \sum_{h=0}^H B_{ss}^h s_s$$

$$+ \sum_{h=0}^H E_{fr}^h \quad (23)$$

$$s.t.(2) - (7), (8a), (17), (22)$$

The problem is then partitioned into a **MP** and a **SP**. The decision variables of **MP** are $X := I_{n,p}^h$, representing production decisions, while the decision variables of **SP** are $Y := (E_{EU}^h, E_{LU}^h)$, corresponding to energy dispatch decisions. This decomposition aligns with the problem's natural structure, where production and energy decisions are treated separately.

$$\begin{aligned} \mathbf{MP} : \min_{I_{n,p}^h, \psi} \sum_{h=0}^H \tilde{T}^h + \psi \\ \text{s.t.} \quad & (2) - (6) \\ \psi &\geq \mathbf{SP}(I_{n,p}^h) \\ & (8a) : \underline{\alpha}_{n,p} - d_{n,p}(I_{n,p}^h) \leq \alpha_{n,p}^h, \quad p \in \mathcal{P} \\ & (22a) : \zeta^h = \sum_i^M \theta_i^h(I_{n,p}^h) \end{aligned}$$

$$\begin{split} \mathbf{SP}(I_{n,p}^{h}) : \max_{\text{DDU, DIU}} \min_{E_{EU}^{h}, E_{LU}^{h}} \sum_{h=0}^{H} E_{fr}^{h} \\ + f_{B}(S^{h}) + \sum_{h=0}^{H} \tilde{P}^{h} \\ + B_{M}^{H} \cdot s_{m} - \zeta^{h} \sum_{h=0}^{H} B_{ss}^{h} \cdot s_{s} \\ \text{s.t.} \quad (7) \\ (17a) : -\Delta E + E_{ex}^{h} \leq \tilde{E}^{h} \leq \Delta E^{h} + E_{ex}^{h} \end{split}$$

This problem is a two-stage robust optimization problem with DDUs. The general form of such problems can be summarized as:

$$\Upsilon \triangleq \{X, Y, W, U, f, G\} \tag{24}$$

where X represents the **MP** decision variable, Y represents the **SP** decision variable, W denotes the DDU set, U denotes the DIU set, f is the objective function, and G is the mapping function linking decision variables to DDUs.

In our formulation for discrete manufacturing, in addition to satisfying the general form outlined in (24), the following two definitions are also satisfied.

Definition 1. (*DDU Independence*) A deterministic DDU variable is associated with only one decision variable, either X or Y.

Definition 2. (DDU Separability) Without loss of generality, a DDU set U is said to be separable if it can be expressed as $U(x) = \{u = C(\xi, x) | \xi \in \Xi\}$. Where $x \in X$ is the decision variable, ξ is an auxiliary random variable, and its uncertainty set Ξ is irrelevant to the decision-making process and is a support set that is independent of the decision variable. $C(\cdot)$: $\Xi \times X \to U$ is a coupling function.

According to Definition 1, the uncertainty set W can be partitioned into two subsets: W_X , representing the decision-dependent uncertainties associated with the master problem

(MP) decision variables X, and W_Y , representing those associated with the sub-problem (\mathbf{SP}) decision variables Y. The mapping functions $G_x: X \to W_X$ and $G_y: Y \to W_Y$ are introduced to describe the relationships between decision variables and their corresponding uncertainties, enabling the reformulation of (24) as (25). The DDU variable separation in Definition 1 is achieved using the linearization method described in Section III.

$$\Upsilon^{\text{DDU}} \triangleq \{X, Y, W_X, W_Y, U, f, G_x, G_y\} \tag{25}$$

In our framework, $I_{n,p}^h$ serves as the **MP** decision variable X, while E_{EU}^h and E_{LU}^h are the **SP** decision variables Y. The DDU variable E_{fr}^h is associated with Y, whereas $(\alpha_{n,p}^h, \zeta^h)$ are linked to X.

Following linearization, the optimization problem in (23) is transformed into a mixed-integer linear programming (MILP) problem, where both the objective functions and constraints of the **MP** and **SP** are expressed in linear form. Given that Ydepends linearly on X and Y, and G_y represents a one-to-one mapping, Y can be treated as a mapping of X and W_Y .

When W_X is empty, the problem with separated DDUs simplifies to a standard two-stage robust optimization problem solvable by traditional algorithms such as Column-and-Constraint Generation (C&CG). However, since $W_X \neq \emptyset$ in our case, traditional methods are not directly applicable. To address this, we propose the Decision-Dependent Columnand-Constraint Generation (DDCCG) algorithm.

The DDCCG algorithm is an iterative approach tailored for decision-dependent two-stage robust optimization problems. The core idea is to iteratively generate and refine constraints for W_X , leveraging Benders decomposition. The algorithmic procedure is outlined in Algorithm 1, with the significance of each step explained below.

- Step 1: Initialize parameters and variables.
- Step 2: Solve the master problem to obtain first-stage decision variables X.
- Step 3: Based on the solution for X, refine the W_X constraints and obtain the relaxed values for W_X .
- Step 4: Solve the sub-problem using the updated X and W_X , obtaining optimal values for Y and W_Y .
- Step 5: Check the termination condition. If not satisfied, generate new constraints (cutting planes) and iterate.

The convergence and computational complexity of the DD-CCG algorithm are analyzed in the next subsection. This algorithm offers fast convergence and adaptability, making it suitable for modern industrial applications.

B. Algorithm Performance Analysis

Under Definitions 1 and 2, our co-scheduling problem is formulated as a two-stage robust optimization model with an additional W_X -type DDU computation layer. In this section, we analyze the convergence and computational complexity of the proposed DDCCG algorithm.

1) Convergence: The convergence of the DDCCG algorithm is established in the following Proposition 1.

Proposition 1. For a decision-dependent two-stage robust optimization problem with both W_X and W_Y types of DDU

Algorithm 1 DDCCG Algorithm

- 1: Set $LB=-\infty$, $UB=+\infty$, k=0, $\mathcal{O}=\emptyset$.
- 2: Solve the master problem.

$$\begin{aligned} \mathbf{MP} &: \min_{I_{n,p}^h, \psi} \sum_{h=0}^H \tilde{T}^h + \psi \\ s.t. \quad &(2) - (6) \\ &\psi \geq \mathbf{SP}(I_{n,p}^h, E_{EU,l}^h, E_{LU,l}^h), \quad \forall l \in \mathcal{O} \\ &\underline{\alpha}_{n,p} - d_{n,p}(I_{n,p}^h) \leq \alpha_{n,p,l}^h, \quad p \in P, \forall l \leq k \\ &\zeta_l^h = \sum_i^M \theta_i^h(I_{n,p}^h), \quad \forall l \leq k \\ &\tilde{E}^h = E^h \cdot e_{n,p} - E_{EU,l}^h - E_{LU,l}^h, \forall l \leq k \end{aligned}$$

Derive an optimal solution $(I_{n,p,k+1}^{h*},\psi_{k+1}^*)$. Update $L\!B\!=\mathbf{MP}(I_{n,p,k+1}^{h*},\psi_{k+1}^*)$. 3: Equate the W_X -type DDU constraint with $I_{n,p,k+1}^{h*}$.

$$\begin{split} \underline{\alpha}_{n,p} - d_{n,p}(I_{n,p,k+1}^{h*}) &= \alpha_{n,p,k+1}^h, \quad p \in P \\ \zeta_{k+1}^h &= \sum_i^M \theta_i^h(I_{n,p,k+1}^{h*}) \end{split}$$

4: Call the oracle to solve $\mathbf{SP}(I_{n,n,k+1}^{h*},\alpha_{n,n,k+1}^h,\zeta_{k+1}^h)$

$$\begin{aligned} \max_{E_{fr}^{h}} \min_{E_{EU}^{h}, E_{LU}^{h}} \sum_{h=0}^{H} E_{fr}^{h} \\ &+ f_{B}(S^{h}) + \sum_{h=0}^{H} \tilde{P}^{h} \\ &+ s_{m} B_{M}^{H} - \zeta^{h} \sum_{h=0}^{H} s_{s} B_{ss}^{h} \end{aligned}$$

s.t.
$$(7a) - (7g), (17a)$$

$$E^{h} = \sum_{n=1}^{N} \sum_{p=1}^{P} I_{n,p,k+1}^{h*} \cdot e_{n,p}$$

$$\tilde{E}^{h} - E(E_{EU}^{h}, E_{LU}^{h}) \leq E_{ex}^{h}$$

Derive $\mathbf{SP}^* = \mathbf{SP}^*(I_{n,p,k+1}^{h*}, \alpha_{n,p,k+1}^h, \zeta_{k+1}^h)$ Update $UB = \min\{UB, LB - \psi_{k+1}^* + \mathbf{SP}^*\}$.

5: if UB- $LB \le \epsilon$, return x_{k+1}^* and terminate. Otherwise, do: (a) if $\mathbf{SP}^* \leq +\infty$, then:

Create $E_{EU,k+1}^h$, $E_{LU,k+1}^h$, add following constraints.

$$\psi \ge \mathbf{SP}(I_{n,p}^h, E_{EU,k+1}^h, E_{LU,k+1}^h)$$
$$\tilde{E}^h = E^h - E_{EU,k+1}^h - E_{LU,k+1}^h$$

to **MP** and update k = k + 1, $\mathcal{O} = \mathcal{O} \cup \{k + 1\}$ and go to Step 2

(b) if $\mathbf{SP}^* = +\infty$, then:

Create $E_{EU,k+1}^h, E_{LU,k+1}^h$, add following constraints.

$$\tilde{E}^h = E^h - E^h_{EU,k+1} - E^h_{LU,k+1}$$

to MP and update k = k + 1 and go to Step 2.

variables satisfying Definitions 1 and 2, the DDCCG algorithm converges to an optimal solution in a finite number of iterations.

Proof. In each iteration, the feasible set of Y is updated based on X_{k+1}^* and $W_{X,k+1}$ derived from **Steps 2–3**. Under Definitions 1 and 2, the feasible set of the MP variable X consists of a finite number of discrete points within a polyhedron, and the feasible set of the \mathbf{SP} variable Y is a

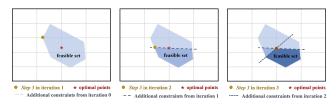


Fig. 3: Constraint generation and convergence in DDCCG algorithm.

union of polyhedra. Given the finiteness of X, the feasible set of Y at optimality reduces to a single polyhedron.

To ensure convergence, we need to demonstrate that the optimal solution for Y in the presence of W_Y -type DDUs can be determined in a finite number of steps. This corresponds to proving the convergence of **Steps 4–5** of the algorithm. In our model, the W_Y -type DDU constraints are expressed as (17a).

The feasible set for Y comprises N_x polyhedra, each defined by constraints such as (17a). For any given polyhedron, we denote the extreme points and extreme rays of $W_Y = E_{fr}^h$ as $W_Y^P \triangleq \{w_{Y1}^P, \ldots, w_{YI}^P\}$ and $W_Y^R \triangleq \{w_{Y1}^R, \ldots, w_{YJ}^R\}$, respectively, where I and J are finite. Extreme rays correspond to feasibility requirements, while extreme points represent conditions for optimality [24]:

$$(E_{fr}^h)_j^R \ge \mathcal{D}(\tilde{E}^h - E_{ex}^h), \quad \forall j = 1, ..., J$$
 (26a)
 $\psi \ge \mathbf{SP}(I_{n,n}^{h*}, (E_{fr}^h)_i^P), \quad \forall i = 1, ... I$ (26b)

In the **MP** stage, the DDCCG algorithm generates candidate solutions $(I_{n,p}^{h*},\alpha_{n,p}^{h*},\zeta^{h*})$ along with a potential optimal value \mathbf{SP}^* represented by ψ . These values are then substituted into the \mathbf{SP} , producing $\mathbf{SP}(I_{n,p}^{h*})$. At this stage, three scenarios may arise:

- Case 1: $SP(I_{n,p}^{h*}) = SP^*$: The algorithm terminates, having found the optimal solution.
- Case 2: $SP(I_{n,p}^{h*}) \neq SP^*$:
 - **Subcase 2.1:** $\mathbf{SP}(I_{n,p}^{h*})$ is unbounded: Additional constraints on W_X -type DDUs are generated and added to the **MP**. These constraints, known as "Benders feasibility cuts," enforce feasibility conditions for the **SP**.
 - Subcase 2.2: $SP(I_{n,p}^{h*}) > SP^*$: In addition to feasibility cuts, "Benders optimality cuts" are introduced, which enforce optimality conditions for the SP.

Since both I and J are finite, the algorithm terminates in a finite number of steps, converging to the optimal solution.

Figure 3 illustrates the convergence process, where the feasible region for W_X is iteratively refined through additional constraints, gradually compressing the solution space until X^* , W_Y^* , and Y^* coincide at the optimal point.

2) *Complexity Analysis*: In this subsection, we provide a precise estimation of the computational complexity of the DDCCG algorithm. First, the optimization problem in this paper satisfies the following assumption:

Assumption 1. The decision-dependent two-stage robust optimization model has the relatively complete recourse property and satisfies the Slater condition.

Under Assumption 1, the following Proposition 2 is therefore put forth.

Proposition 2. Consider a decision-dependent two-stage robust optimization problem. Let N_x represent the number of discrete feasible points in the decision space of X, and let N_{W_Y} denote the number of extreme points of the polyhedral set W_Y associated with **SP** under a fixed X. Under these conditions, the DDCCG algorithm is guaranteed to converge to the optimal solution within $O(N_x N_{W_Y})$ iterations.

Proof. The value of W_X is determined by the feasible set of X, which consists of N_x discrete points. Proving Proposition 2 reduces to demonstrating that the complexity of solving a two-stage robust optimization problem without W_X (i.e., by omitting **Step 3**) is $O(N_{W_Y})$.

We rewrite our decision-dependent two-stage robust optimization model with only W_Y under a fixed W_X in a general form. Using the aforementioned linearization techniques, W_Y is decomposed into a component associated with Y and a DIU variable \mathcal{U} , where \mathcal{U} is a polyhedron. The resulting problem is formulated as follows.

$$\min_{x} c^{T} x + \max_{u \in \mathcal{U}} \min_{y} b^{T} y$$
s.t. $Ax \ge d$

$$Gy \le Ex$$

$$By \le e$$

$$Hy \ge k - Mu$$

$$x \in S_{x}, y \in S_{y}$$

$$(27)$$

Under Assumption 1, the problem can be converted into the dual form by invoking the strong duality property:

$$\min_{x} c^{T} x + \max_{u,\lambda,\gamma,\mu} e^{T} \gamma + (k - Mu)^{T} \mu + y^{T} E^{T} \lambda$$
s.t. $Ax \ge d$

$$G^{T} \lambda + B^{T} \gamma + H^{T} \mu = b$$

$$\lambda, \gamma, \mu \ge 0$$

$$u \in \mathcal{U}, y \in S_{y}$$

$$(28)$$

Note that the inner maximization problem is a bilinear program over two disjoint polyhedra, its optimal solution always corresponds to extreme points of these polyhedra. Let $\hat{\mathcal{U}} = \{u_1, \dots, u_{N_{W_Y}}\}$ represent the collection of extreme points of \mathcal{U} . Thus, the objective function of the original problem 27 is equivalent to:

$$\min_{x} c^{T} x + \max_{u \in \hat{\mathcal{U}}} \min_{y} b^{T} y \tag{29}$$

Enumerating values in the finite set \hat{U} , the problem can be reformulated as a mixed-integer program:

$$\min_{x} c^{T} x + \eta$$
s.t. $Ax \ge d$

$$\eta \ge b^{T} y_{l}, \quad l = 1, ..., N_{W_{Y}}$$

$$Gy_{l} \leq Ex, \quad l = 1, ..., N_{W_{Y}}$$
 (30)
 $By_{l} \leq e, \quad l = 1, ..., N_{W_{Y}}$
 $Hy_{l} \geq k - Mu_{l}, \quad l = 1, ..., N_{W_{Y}}$
 $x \in S_{x}, y_{l} \in S_{y}, \quad l = 1, ..., N_{W_{Y}}$

For this formulation of a decision-dependent two-stage robust optimization model, we denote the **MP** and **SP** in our DDCCG algorithm as \mathbf{MP}_0 and \mathbf{SP}_0 for clarity. The \mathbf{MP}_0 is a relaxation providing a lower bound for the problem (30).

$$\mathbf{MP_0} : \min_{x,\eta} c^T x + \eta$$
s.t. $Ax \ge d$,
$$\eta \ge b^T y_l, \quad \forall l \in \mathcal{O},$$

$$Gy_l \le Ex, \quad \forall l \le k,$$

$$By_l \le e, \quad \forall l \le k,$$

$$Hy_l \ge k - Mu_l, \quad \forall l \le k,$$

$$x \in S_x, \ y_l \in S_y, \quad \forall l \le k.$$
(31)

$$\begin{aligned} \mathbf{SP}_0 : & \max_{u \in \mathcal{U}} & \min_{y} b^T y \\ \text{s.t.} & & Gy \le Ex, \\ & & By \le e, \\ & & Hy \ge k - Mu, \\ & & y \in S_y \end{aligned} \tag{32}$$

Note that solving SP_0 expands MP_0 by including an additional u_l along with its corresponding recourse variables and constraints. We now demonstrate that any repeated u^* during the procedure implies optimality, i.e., LB = UB in the DDCCG algorithm. Assume that in the k-th iteration, (x^*, η^*) and (u^*, y^*) are the optimal solutions to \mathbf{MP}_0 and \mathbf{SP}_0 of the primal problem, respectively, and that u^* has already been identified in a previous iteration. From Step 4 of the DDCCG algorithm, we know: $UB \le c^T x^* + b^T y^*$. Since u^* has been identified in a prior iteration, the current \mathbf{MP}_0 is identical to the **MP**₀ from iteration (k-1). Therefore, (x^*, η^*) must also be the optimal solution of \mathbf{MP}_0 in the (k-1)-th iteration. From Step 2 of the DDCCG algorithm, we have: $LB \geq c^T x^* + \eta^* \geq c^T x^* + b^T y^*$, where the final inequality follows from the fact that u^* was already identified and the associated constraints were added to MP_0 in or before the (k-1)-th iteration. Consequently, LB = UB, and the optimality condition is satisfied. The conclusion follows directly from the finiteness of \mathcal{U} , the set of all extreme points of the polyhedral uncertainty set. Thus, the complexity of solving the two-stage robust optimization problem without W_X is $O(N_{W_Y})$.

Remark 4. Compared to Proposition 1, Proposition 2 provides a precise complexity estimate for problems satisfying Assumption 1.

V. CASE STUDY

This section begins with the modeling of a real-world engine assembly factory using the Petri net framework. The proposed model and the DDCCG algorithm are subsequently employed to generate simulation results, which are further analyzed and discussed.

A. Petri Net Modeling

To validate the efficacy of the proposed methodology, we utilize a real-world engine assembly factory as a representative case study for discrete manufacturing [25]. The Petri net framework is adopted for its capability to quantitatively describe complex industrial processes [26]. Petri nets are particularly effective in characterizing the sequential and concurrent relationships between events and operations over a temporal sequence [26]. The framework comprises two fundamental components: *places*, representing system states, and *transitions*, representing state transitions. The production, assembly, transportation, and testing stages of the engine assembly process are modeled using a Petri net (see Fig 8 in the Appendix).

The entire industrial process, spanning from raw materials to final products (engines), is represented through 12 workshops, 14 types of equipment, and multiple buffers (node labels detailed in Table III in the Appendix). Based on the specific functions of each workshop, the production process is categorized into eight major segments:

- Parts Production: stamping machine (SM).
- **Datum Milling**: milling machine, face milling machine.
- Crankshaft Grinding: circular grinding machine (CG), CNC grinding machine.
- Crankshaft Mounting: automatic assembly line (AAL₁), robotic assembly line (RAL₁).
- Piston Assembly: piston assembly machine (PAM), automatic assembly machine (AAM₁).
- Cylinder Mounting: automatic assembly line (AAL₂), robotic assembly line (RAL₂).
- Toothed Belt Mounting: belt mounting machine (BMM), automatic assembly machine (AAM₂).
- **Testing**: integrated test platform.

In real-world production scenarios, the sequence of certain stages, such as crankshaft grinding and cylinder mounting, can be adjusted. Additionally, if a semi-finished product undergoes reprocessing on the same equipment, a detector will inspect or repair it before directing it to the buffer. This reprocessing step enhances the production yield. Moreover, the P54 node is designated as the by-product output buffer, handling by-products such as cylinder liners. The simulation assumes a continuous 24-hour production cycle, replicating real-world operating conditions.

B. Analysis of Results

1) Reduced Operational Cost: Our proposed method demonstrates substantial potential in reducing production costs. As shown in Table I, our model, whether implemented without DDUs or with DDUs, consistently outperforms both historical and ideal data in terms of operational cost. Unlike conventional factory operations, the model enables integrated scheduling of energy and production equipment, leading to reduced power consumption and increased overall production

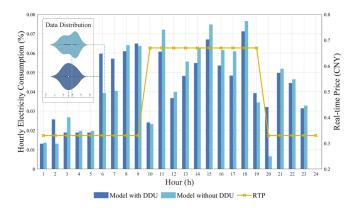


Fig. 4: Electricity consumption comparison under RTP.

within the industrial park. Furthermore, compared to the model without DDUs, the DDU-enabled model achieves comparable main product output while simultaneously reducing energy consumption and increasing by-product production.

The reduction in electricity costs achieved by the DDU-enabled model relative to the model without DDUs is primarily due to its ability to effectively avoid electricity consumption during peak periods. As illustrated in Fig 4, during the RTP peak periods (i.e., when the golden folded line reaches its peak plateau), the DDU-enabled model (dark blue) exhibits a significant decrease in electricity consumption compared to the model without DDUs (light blue), with reductions of approximately 3–5% for each segment of the bar chart. This highlights the model's capability to dynamically adjust energy usage based on real-time price signals, contributing to cost efficiency.

In addition to reducing energy consumption, the increase in by-product production observed in the DDU-enabled model compared to the model without DDUs can be attributed to the decentralization of production. This decentralization facilitates the optimal utilization of redundant raw materials from primary product manufacturing. The total utilization frequency for each device over 24 hours is calculated and visualized in a matrix, with device locations considered. Bilinear interpolation is applied to expand the matrix, generating a contour map as depicted in Fig 5.

Compared to the model without DDUs, the DDU-enabled model forms two distinct production centers of gravity (highlighted in yellow areas). Notably, the production line on the right contains the P54 node, designated for the storage and sale of by-products. This indicates that leveraging redundant materials for by-product synthesis and subsequent sales is an effective strategy for enhancing profitability. Additionally, the establishment of a multi-center production strategy increases the resilience of the production system. In the event of a disruption to one production line, the operational load can be redistributed to another center of gravity, thereby mitigating the impact of such disruptions.

Overall, the integrated scheduling and decision-dependent modeling approach effectively balances energy efficiency and production flexibility, offering a comprehensive solution for modern industrial operations to reduce costs.

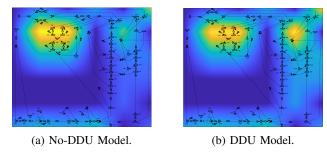


Fig. 5: Visualization of decentralized production effects.

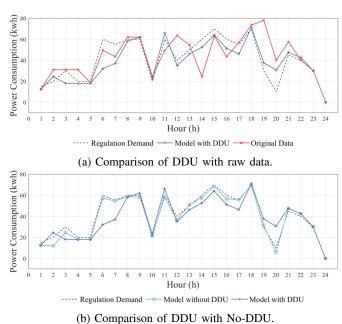


Fig. 6: Comparison of FR realizations across different methods

2) Frequency Regulation & Peak Shaving: The proposed model effectively adapts to frequency regulation demands and enhances the system's peak shaving capabilities. Fig 6 presents the results of a comparative analysis among the DDU-enabled model, the model without DDUs, and the original dataset, in response to specified frequency regulation requirements. Both the DDU and non-DDU models align more closely with the frequency regulation demand curve than the original dataset. However, the DDU model demonstrates a less aggressive response to frequency regulation, exhibiting a smoother power consumption curve (i.e., improved peak shaving performance).

The DDU model achieves superior peak shaving performance by smoothing the production curve compared to the non-DDU model. The upper left part of Fig 4 illustrates violin plots of electricity usage for the two models, highlighting the DDU-enabled model's smoother curve and smaller variance. This performance stems from the DDU model's consideration of higher uncertainty, prompting the system to adopt a more conservative and stable power usage strategy to address potential fluctuations, such as RTP spikes. From the power grid perspective, this behavior corresponds to increased usage of BESS storage and advanced power planning strategies. The

TABLE I: Comparison of operational costs across different methods.

ModelType	Power Cost (CNY)	Average Cost (CNY)	Main Products (CNY)	By-products (CNY)	Objective Function (CNY)
Original Data	1042.00	43.42	2250	200	-
Expected Data	980.00	40.83	2250	200	-
Model without DDU	945.09	39.39	2500	160	-2936.41
*Model with DDU	922.97	38.46	2500	250	-3340.78

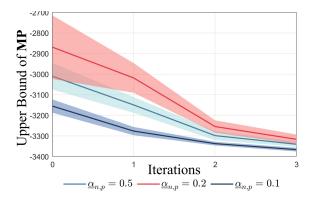


Fig. 7: Uncertainty error in the upper bound of optimization problem under different $\underline{\alpha}_{n,p}$ values.

reduced equipment operation during peak RTP periods further underscores this approach.

Thus, the proposed DDU model effectively balances frequency regulation performance and peak shaving capabilities, offering significant industrial and practical benefits by simultaneously addressing energy cost and system resilience.

3) **DDU Impact Assessment**: In evaluating the impact of introducing DDUs, we analyze the performance of the proposed model under different scenarios.

For the product yield DDU, the yield inherent in the equipment, represented by $\underline{\alpha}_{n,p}$, influences the model's estimation of uncertainty. Specifically, $\underline{\alpha}_{n,p}$ quantifies the intrinsic uncertainty associated with the device's yield. Fig 7 illustrates the model's upper bound (the optimal value of the master problem) over iterations for varying values of $\underline{\alpha}_{n,p}$. The error band in Fig 7 reflects the relative interval width for S_1 in the **MP**. As the **SP** iteratively adds constraints, the error band narrows, indicating a preference for production line combinations with reduced uncertainty intervals. Notably, for $\underline{\alpha}_{n,p} = 0.1$, the gap does not significantly shrink during later iterations. This suggests that yield is not the sole determinant of the final profit. Furthermore, constraints from the SP are not always fully satisfied for S_1 in the **MP**, highlighting inherent modeling limitations. Models with greater $\underline{\alpha}_{n,p}$ values demonstrate slower convergence, inferior optimal solutions, and larger error bands. This behavior arises from the increased difficulty in estimating DDUs due to substantial intrinsic yield errors, which widen the error band and reduce yield. Consequently, a larger $\underline{\alpha}_{n,n}$ results in diminished model performance.

For the frequency regulation penalty DDU, the plant operator dynamically adjusts the target weights based on anticipated benefits. We found that compared to the model without DDUs, the model with DDUs performs worse in frequency regulation.

TABLE II: Optimal values under different γ

γ	By-products Tendency	Objective Function (CNY)
0.01	$\downarrow \downarrow$	-3319.52
0.02	↓	-3328.47
0.05	†	-3340.78
0.10	$\uparrow\uparrow$	-3312.27

This shift reflects the operator's focus on transitioning from frequency regulation to production objectives. The rate of this transition depends on the degree of uncertainty in the optimization problem. When the uncertainty variable is associated with decisions, it causes the model to face more uncertainties and struggle to accurately estimate expected frequency regulation requirements.

For the product structure DDU, a local optimum is observed in the ratio of primary to secondary products. To examine this effect, we vary the confidence level γ , which directly adjusts the values of ζ^h , $\underline{\theta}^h$, and $\bar{\theta}^h$. A larger γ increases both the upper and lower bounds of the by-product share, thereby expanding the range of possible by-product outputs. Table II summarizes the optimal values of the objective function for different γ levels. The results indicate that the objective value initially decreases and subsequently increases as the ratio of primary to by-products decreases. This behavior occurs because moderate by-product production effectively utilizes redundant raw materials, while excessive by-product output, due to its lower profitability, reduces overall profits.

In summary, the introduction of DDUs allows for more nuanced decision-making by accounting for uncertainties inherent in yield, FR penalties, and product structures. This capability enables the model to achieve better trade-offs between production efficiency, energy usage, and profitability, demonstrating its potential for practical industrial applications.

VI. CONCLUSION

The implementation of a co-scheduling strategy for energy and production is critical in discrete manufacturing systems. In this paper, we propose a two-stage robust optimization model incorporating DDUs to address energy management and equipment dispatch challenges. Unlike traditional models that focus solely on DIUs, the proposed model accounts for multiple types of DDUs, including ambiguity sets and probability distributions. The proposed model is constructed and linearized to facilitate tractable optimization, and a novel algorithm, named DDCCG, is developed to efficiently solve the decision-dependent two-stage robust optimization problem. The algorithm demonstrates rapid convergence and provides

optimal solutions for both the operational points of production equipment and the energy access strategies of the energy system. To validate the effectiveness of the proposed approach, we conducted a case study using a real-world engine assembly line modeled with the Petri net method. The results highlight several advantages of the model: (1) it effectively reduces production costs by minimizing electricity demand without compromising the primary product output, (2) it achieves superior frequency regulation performance and enhances peak shaving capabilities, and (3) it improves the resilience of the production line by enabling the system to adapt to risks and uncertainties more effectively. In conclusion, the findings of this study demonstrate the significant potential of the proposed co-scheduling strategy in enhancing production performances. The integration of DDU considerations and the DDCCG algorithm offers a robust framework for optimizing production and energy operations in complex industrial environments. The proposed method has promising applications for reducing operational costs and improving system resilience in future industrial settings.

APPENDIX A PETRI NET NODE DESCRIPTION

Engine assembly line Petri net nodes are described as Fig 8 and Table III.

REFERENCES

- [1] H. Wang and W. Chen, "Modelling deep decarbonization of industrial energy consumption under 2-degree target: Comparing china, india and western europe," *Applied Energy*, vol. 238, pp. 1563–1572, 2019.
- [2] G. May, B. Stahl, M. Taisch, and D. Kiritsis, "Energy management in manufacturing: From literature review to a conceptual framework," *Journal of cleaner production*, vol. 167, pp. 1464–1489, 2017.
- [3] M. Parente, G. Figueira, P. Amorim, and A. Marques, "Production scheduling in the context of industry 4.0: review and trends," *Interna*tional Journal of Production Research, vol. 58, no. 17, pp. 5401–5431, 2020.
- [4] A. W. Dowling and V. M. Zavala, "Economic opportunities for industrial systems from frequency regulation markets," *Computers & Chemical Engineering*, vol. 114, pp. 254–264, 2018.
- [5] T. Kern and B. Bukhari, "Peak shaving—a cost-benefit analysis for different industries," Proceedings of the Internationale Energiewirtschaftstagung (IEWT), Vienna, Austria, pp. 8–10, 2021.
- [6] M. Ghaleb, H. Zolfagharinia, and S. Taghipour, "Real-time production scheduling in the industry-4.0 context: Addressing uncertainties in job arrivals and machine breakdowns," *Computers & Operations Research*, vol. 123, p. 105031, 2020.
- [7] J. Englberger, F. Herrmann, and M. Manitz, "Two-stage stochastic master production scheduling under demand uncertainty in a rolling planning environment," *International journal of production research*, vol. 54, no. 20, pp. 6192–6215, 2016.
- [8] D. Tang, M. Dai, M. A. Salido, and A. Giret, "Energy-efficient dynamic scheduling for a flexible flow shop using an improved particle swarm optimization," *Computers in industry*, vol. 81, pp. 82–95, 2016.
- [9] M. Ghaleb, H. Zolfagharinia, and S. Taghipour, "Real-time production scheduling in the industry-4.0 context: Addressing uncertainties in job arrivals and machine breakdowns," *Computers & Operations Research*, vol. 123, p. 105031, 2020.
- [10] W. Qian, Y. Guo, H. Zhang, S. Huang, L. Zhang, H. Zhou, W. Fang, and S. Zha, "Digital twin driven production progress prediction for discrete manufacturing workshop," *Robotics and Computer-Integrated Manufacturing*, vol. 80, p. 102456, 2023.
- [11] J. Dupacová, "Optimization under exogenous and endogenous uncertainty," University of West Bohemia in Pilsen, 2006.
- [12] Y. Zhang, F. Liu, Y. Su, Y. Chen, Z. Wang, and J. P. Catalão, "Two-stage robust optimization under decision dependent uncertainty," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 7, pp. 1295–1306, 2022.

- [13] N. H. Lappas and C. E. Gounaris, "Robust optimization for decision-making under endogenous uncertainty," *Computers & Chemical Engineering*, vol. 111, pp. 252–266, 2018.
- [14] A. D. Bonzanini, J. A. Paulson, and A. Mesbah, "Safe learning-based model predictive control under state-and input-dependent uncertainty using scenario trees," in 2020 59th IEEE Conference on Decision and Control (CDC), pp. 2448–2454, IEEE, 2020.
- [15] B. Hu, C. Pan, C. Shao, K. Xie, T. Niu, C. Li, and L. Peng, "Decision-dependent uncertainty modeling in power system operational reliability evaluations," *IEEE Transactions on Power Systems*, vol. 36, no. 6, pp. 5708–5721, 2021.
- [16] R. V. Yohanandhan, R. M. Elavarasan, P. Manoharan, and L. Mihet-Popa, "Cyber-physical power system (cpps): A review on modeling, simulation, and analysis with cyber security applications," *IEEE Access*, vol. 8, pp. 151019–151064, 2020.
- [17] J. Forman, J. Stein, and H. Fathy, "Optimization of dynamic battery paramter characterization experiments via differential evolution," in 2013 American Control Conference, pp. 867–874, IEEE, 2013.
- [18] Z. Ma, S. Zou, and X. Liu, "A distributed charging coordination for large-scale plug-in electric vehicles considering battery degradation cost," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 2044–2052, 2015.
- [19] Y. Shi, B. Xu, D. Wang, and B. Zhang, "Using battery storage for peak shaving and frequency regulation: Joint optimization for superlinear gains," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2882– 2894, 2017.
- [20] S. Hao and A. Papalexopoulos, "External network modeling for optimal power flow applications," *IEEE transactions on power systems*, vol. 10, no. 2, pp. 825–837, 1995.
- [21] Y. Zhang, R. Jiang, and S. Shen, "Ambiguous chance-constrained binary programs under mean-covariance information," SIAM Journal on Optimization, vol. 28, no. 4, pp. 2922–2944, 2018.
- [22] P. Walley, Statistical reasoning with imprecise probabilities, vol. 42. Springer, 1991.
- [23] Y. Zhang, X. Han, M. Yang, B. Xu, Y. Zhao, and H. Zhai, "Adaptive robust unit commitment considering distributional uncertainty," *Interna*tional Journal of Electrical Power & Energy Systems, vol. 104, pp. 635– 644, 2019.
- [24] Z. C. Taskın, "Benders decomposition," Wiley Encyclopedia of Operations Research and Management Science. John Wiley & Sons, Malden (MA), 2010.
- [25] A. Jafari Asl, M. Solimanpur, and R. Shankar, "Multi-objective multi-model assembly line balancing problem: a quantitative study in engine manufacturing industry," *Opsearch*, vol. 56, pp. 603–627, 2019.
- [26] J. L. Peterson, "Petri nets," ACM Computing Surveys (CSUR), vol. 9, no. 3, pp. 223–252, 1977.

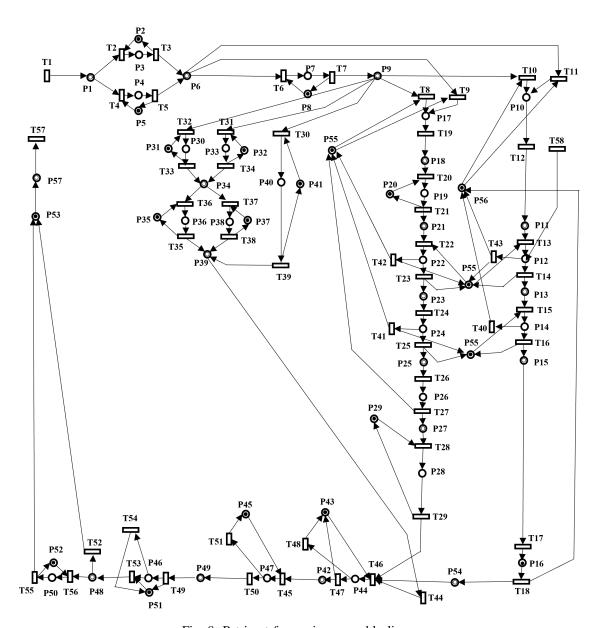


Fig. 8: Petri net for engine assembly lines.

TABLE III: Petri net: places and transitions.

Places		Transitions		
P1: Buffer0	P30: Crankshaft on CG	T1: Raw material input	T30: Production information input	
P2: SM	P31: CG	T2: SM start	T31: CNC grinder start	
P3: Raw material in SM	P32: CNC grinder	T3: SM running	T32: CG start	
P4: Raw material in SM	P33: Crankshaft on CNC grinder	T4: SM start	T33: CG running	
P5: SM	P34: Buffer3	T5: SM running	T34: CNC grinder running	
P6: Buffer1	P35: RAL ₂	T6: Millin, face milling start	T35: RAL ₂ running	
P7: Parts in milling, face milling	P36: Cylinder in RAL ₂	T7: Millin, face milling running	T36: RAL ₂ start	
P8: Millin, face milling	P37: RAL ₂	T8: CNC grinder feed system start	T37: RAL ₂ start	
P9: Buffer2	P38: Cylinder in RAL ₂	T9: CNC grinder start	T38: RAL ₂ running	
P10: Crankshaft on CG	P39: Buffer8	T10: CG start	T39: Information processing	
P11: Buffer4-1	P40: Information control system	T11: CG start	T40: Sampling and transportation	
P12: Main bearing in AAL ₁	P41: Information control system	T12: CG running	T41: Sampling and transportation	
P13: Buffer5-1	P42: Buffer9	T13: AAL ₁ start	T42: Sampling and transportation	
P14: Semi-finished in machine	P43: AAL ₁	T14: AAL ₁ running	T43: Sampling and transportation	
P15: Buffer6-1	P44: Main bearing in AAL ₁	T15: PAM start	T44: Cargo transportation	
P16: Cylinder in AAL ₂	P45: AAM ₂	T16: PAM running	T45: AAM ₁ start	
P17: Crankshaft on CNC grinder	P46: Belt on AAM ₂	T17: AAL ₂ start	T46: AAL ₁ start	
P18: Buffer4-2	P47: Semi-finished in AAM ₂	T18: AAL ₂ running	T47: AAL ₁ running	
P19: Crankshaft on CNC grinder	P48: Buffer11	T19: CNC grinder	T48: AAL ₁ running	
P20: Engineer communication unit	P49: Buffer10	T20: CNC grinder running	T49: BMM, AAM ₂ start	
P21: Buffer4-3	P50: Main product in IDU	T21: Bearing output	T50: AAM ₁ running	
P22: Bearing in RAL ₁	P51: BMM, AAM ₂	T22: RAL ₁ start	T51: AAM ₁ running	
P23: Buffer5-2	P52: AI detection unit	T23: RAL ₁ running	T52: Sampling and transportation	
P24: Semi-finished in machine	P53: Integrated test platform	T24: PAM start	T53: BMM, AAM ₂ running	
P25: Buffer6-2: Register	P54: Buffer7-2	T25: PAM running	T54: BMM, AAM2 running	
P26: Cylinder in AAL ₂	P55: Worker inspection station	T26: AAL ₂ start	T55: AI detection unit running	
P27: Buffer7-1	P56: Worker inspection station	T27: AAL ₂ running	T56: AI detection unit start	
P28: Defective in repair station	P57: Buffer12	T28: Defective repair	T57: Main product output	
P29: Engineer repair station		T29: Defective transportation	T58: Main bearing input	