Integrated Scheduling with Equipment Health, Maintenance, and Retroactive Quality Control in Semiconductor Manufacturing with Fixed Sequence

Mourad Terzi [1] and Claude Yugma[1]

Abstract— **This paper investigates an integrated scheduling problem on a single machine in a semiconductor manufacturing environment, where equipment degradation and product quality are jointly considered. Equipment degradation, monitored through an Equipment Health Indicator (EHI), directly impacts both production feasibility and quality. The presented study focuses on the case where the job sequence is fixed in advance, a situation frequently encountered in practice due to technological or operational constraints. Preventive maintenance, retroactive quality sampling, and job rejection are jointly optimized, with the objective of minimizing the weighted sum of job tardiness and operational costs. The problem is formulated as a Mixed-Integer Linear Programming ($MILP$) model that incorporates equipment degradation, sequencing, maintenance, and retroactive sampling constraints. Theoretical properties related to the problem's *NP-hardness* are established. Finally, computational experiments on randomly generated semiconductor manufacturing instances are performed to analyse the model's behavior and evaluate its performance under various conditions.**

I. INTRODUCTION AND RELATED WORKS

Semiconductor manufacturing is well known for its high level of complexity, both in terms of production flow and operational constraints. Modern fabrication facilities involve expensive equipment, re-entrant flows, and strict quality requirements, making scheduling decisions a key factor in throughput, cost, and delivery performance. These issues have been widely discussed in the literature, for instance in the survey by [10], which highlighted the specific difficulties of scheduling in semiconductor fabs compared to classical manufacturing environments.

Most scheduling models rely on the simplifying assumption that machine performance remains constant over time. In practice, this assumption is rarely satisfied. Processing tools progressively degrade as they are used, and this degradation may affect the capability of the equipment to produce conforming products. As a consequence, decisions that are optimal from a purely scheduling perspective may lead to undesirable situations, such as producing a large number of lots under degraded conditions and discovering quality problems only afterwards. In industrial environments where the value of a lot can be very high, such situations are particularly critical.

With the development of Advanced Process Control (APC) and Industry 4.0 technologies, it has become possible to monitor equipment condition in real time using large amounts of sensor data. These data are often aggregated into health indicators that provide a synthetic view of the machine state. Indicators of this type, sometimes referred to as Equipment Health Indicators (EHI), are increasingly used in practice to detect drifts and anticipate failures [7]. At the same time, quality control in semiconductor manufacturing frequently relies on sampling procedures, where the inspection of one lot may allow the validation of several previously processed lots, provided that no significant disturbance has occurred in the meantime [9], [12]. This creates a temporal dependency between production, maintenance, and inspection decisions: maintenance restores the machine condition, sampling validates past production, and the health indicator determines whether future processing should be allowed.

The interaction between production scheduling and maintenance has been studied for many years [13]. Several authors have shown that optimizing production and maintenance separately may lead to suboptimal results, and that both decisions should be considered jointly [1], [3]. In these works, machine degradation is usually modeled through reliability functions, age-dependent failure rates, or usage-based indicators, and maintenance is inserted in the schedule in order to limit the risk of breakdowns or performance loss. More recent approaches consider condition-based or health-based maintenance, where the decision to perform maintenance depends on the estimated state of the equipment rather than on fixed time intervals [2]. Although these models capture an important aspect of real systems, they generally do not include explicit quality validation mechanisms.

Another line of research focuses on the relationship between production and quality control. Adaptive inspection policies have been proposed in order to adjust the position or frequency of inspections according to the estimated state of the process [4]. In semiconductor manufacturing, metrology data and inline process control are commonly used to detect process drifts and to decide when additional measurements are required [5]. These approaches make it possible to reduce inspection costs while maintaining product quality, but they usually treat inspection as a reactive decision coupled to production, without considering its interaction with maintenance planning.

More recently, several works have explored the use of real-time indicators to support proactive decision-making in man-

*This work was not supported by any organization

[1] Mines Saint-Etienne, Department of Manufacturing Sciences and Logistics, UMR 6158 LIMOS, CMP, Univ. Clermont Auvergne, CNRS, Gardanne, F-13541, France,

ufacturing systems. In these approaches, information about the current or predicted state of the equipment can be used to advance maintenance, delay production, or increase inspection when a risk is detected [8]. Such methods represent an important step toward more integrated decision-making. Nevertheless, in most of the existing literature, scheduling, maintenance, and quality control are still studied either two by two or in a sequential manner. A growing but limited number of contributions propose integrated models where production scheduling, maintenance planning, and quality-related decisions are optimized jointly in a single framework, often showing substantial cost savings compared with decoupled approaches [11]. Such fully integrated models remain relatively rare, especially in settings where equipment degradation directly affects product conformity and thus constrains the feasibility of production, despite some recent reliability-based and quality-centred formulations explicitly linking machine condition, production plans, and quality losses.

This paper studies an integrated decision problem arising in a context frequently observed in semiconductor manufacturing, namely when the processing sequence of jobs is fixed in advance. Such a situation occurs naturally when the order of lots is imposed by recipe constraints, batching decisions, or higher-level planning rules [10]. Even when the sequence cannot be modified, some flexibility remains. A preventive maintenance can be inserted, sampling operations can be activated, and some jobs can be rejected in order to avoid processing them under poor equipment conditions. These decisions are strongly interdependent, because equipment degradation, maintenance, and sampling all evolve along the same time axis.

The objective of this work is to optimize these decisions jointly. A new problem is introduced in which the equipment condition is represented by the EHI, maintenance actions restore the health level, and retroactive sampling can validate previously processed jobs within a limited window. Even with a fixed processing order, the integration of these mechanisms leads to a non-trivial combinatorial problem. The main contributions of this paper are as follows: (i) An integrated scheduling problem combining a fixed job sequence, preventive maintenance, equipment health dynamics, and retroactive sampling is formally defined. (ii) The problem is proven to be NP-hard even when the job order is fixed. (iii) Computational experiments are reported on realistic instances to evaluate the model's performance.

The remainder of the paper is organized as follows. Section II describes the problem and the modelling assumptions. Section III presents the complexity analysis and the mathematical formulation. Section IV reports the numerical experiments, and Section V concludes the paper.

II. PROBLEM DESCRIPTION AND FORMULATION

The problem investigated in this study considers a single machine processing a set of n jobs in a fixed and predetermined sequence. Such a constraint naturally occurs in semiconductor manufacturing environments due to recipe-dependent tool constraints, batching decisions, or master production planning requirements. The machine condition is monitored through an EHI indicator, which is initially at its maximum level EHI_{max} and progressively deteriorates after the processing of each job.

Although the job sequence is fixed, the decision-maker retains several degrees of freedom: (i) inserting preventive maintenance, (ii) rejecting selected jobs, and (iii) activating retroactive quality inspections (sampling). Preventive maintenance fully restores the machine to its initial condition (EHI_{max}). Rejected jobs are omitted from both the temporal schedule and the machine degradation dynamics, which can help limit system deterioration. The retroactive quality control (sampling) validates the K most recent jobs, under the condition that no maintenance has occurred within this backward window of size K. These decisions are interdependent through timing constraints, equipment degradation, and quality validation mechanisms. The overall objective is to minimize the total job tardiness along with the associated operational costs, including: (i) maintenance costs, (ii) sampling costs, (iii) rejection costs, and (iv) non-validation costs. The problem is referred to in the rest of the paper as Fixed-Sequence, Sampling, and Maintenance with K-window quality validation (*FS-SM-K*).

The notations used in the proposed model, including sets, parameters, and decision variables, are summarized below.

A. Notations: sets

- $\sigma = (1, 2, \dots, n)$: fixed sequence of n jobs.
- $t \in \{i, \dots, i + K - 1\}$: sampling position associated with job i.
- $k \in \{i, \dots, t\}$: intermediate positions.

B. Notations: parameters

- p_i: processing time of job i.
- r_i: release date of job i.
- d_i: due date of job i.
- w_i: weight of job i.
- τ_m: duration of a maintenance intervention.
- $EHI_{\max}$: maximum equipment health index.
- δ: degradation rate by unit of processing time.
- θ: critical threshold of the EHI below which production becomes infeasible.
- K: size of the backward sampling window.
- C_m: maintenance cost.
- C_s: sampling cost.
- C_r: rejection cost.
- C_{unv}: cost of an invalidated job.
- M: sufficiently large constant used for constraint.

C. Notations: decision variables

- $\rho_i \in \{0, 1\}$: equals 1 if job i is rejected; 0 otherwise.
- $m_i \in \{0, 1\}$: equals 1 if a maintenance operation is performed before job i; 0 otherwise.
- $z_i \in \{0, 1\}$: equals 1 if job i is either rejected or validated by at least one sampling operation; 0 otherwise.

- $sm_i \in \{0,1\}$: equals 1 if a retroactive sampling is performed immediately after job i; 0 otherwise.
- $y_{it} \in \{0,1\}$: equals 1 if the sampling at position t validates job i; 0 otherwise.
- S_i: start time of job i.
- C_i: completion time of job i.
- C_i^{eff}: effective completion time of job i.
- H_i: equipment health index immediately before processing job i.
- T_i: tardiness of job i.

D. Mathematical formulation

Based on the notation introduced above, the *MILP* formulation of the *FS-SM-K* problem, denoted as P_1, is presented as follows:

$$\min \sum_{i=1}^{n} w_i\big(T_i + C_r\rho_i\big) + C_m m_i + C_s sm_i + C_{unv}(1 - z_i) \tag{1}$$

$$\text{s.t. } S_1 = \tau_m m_1 \tag{2}$$

$$S_{i+1} = S_i + p_i(1 - \rho_i) + \tau_m m_{i+1}, \quad i = 1, \dots, n-1 \tag{3}$$

$$r_i - M\rho_i \le S_i, \quad i = 1, \dots, n \tag{4}$$

$$C_i = S_i + p_i, \quad i = 1, \dots, n \tag{5}$$

$$C_i^{eff} \ge C_i - M\rho_i, \quad i = 1, \dots, n \tag{6}$$

$$C_i^{eff} \le C_i + M\rho_i \quad i = 1, \dots, n \tag{7}$$

$$C_i^{eff} \le M(1 - \rho_i), \quad i = 1, \dots, n \tag{8}$$

$$T_i \ge C_i^{eff} - d_i, \quad i = 1, \dots, n \tag{9}$$

$$H_1 = EHI_{max} \tag{10}$$

$$H_i \le H_{i-1} - \delta p_{i-1}(1 - \rho_{i-1}) + Mm_i, \quad i = 2, \dots, n \tag{11}$$

$$H_i \ge H_{i-1} - \delta p_{i-1}(1 - \rho_{i-1}) - Mm_i, \quad i = 2, \dots, n \tag{12}$$

$$H_i \ge EHI_{max} - M(1 - m_i), \quad i = 1, \dots, n \tag{13}$$

$$H_i \le EHI_{max} + M(1 - m_i), \quad i = 1, \dots, n \tag{14}$$

$$H_i \ge \theta - M\rho_i, \quad i = 1, \dots, n \tag{15}$$

$$m_i \le 1 - \rho_i, \quad i = 1, \dots, n \tag{16}$$

$$sm_i \le 1 - \rho_i, \quad i = 1, \dots, n \tag{17}$$

$$y_{i,t} \le sm_t, \quad i = 1, \dots, n,\ t = i, \dots, \min(i + K - 1, n) \tag{18}$$

$$y_{i,t} \le 1 - \rho_i, \quad i = 1, \dots, n,\ t = i, \dots, \min(i + K - 1, n) \tag{19}$$

$$\begin{aligned} &y_{i,t} \le 1 - m_k, \quad i = 1, \dots, n, \\ &t = i, \dots, \min(i + K - 1, n),\ k = i + 1, \dots, t \end{aligned} \tag{20}$$

$$z_i \ge \rho_i, \quad i = 1, \dots, n \tag{21}$$

$$z_i \ge y_{i,t}, \quad i = 1, \dots, n,\ t = i, \dots, \min(i + K - 1, n) \tag{22}$$

$$z_i \le \rho_i + \sum_{t=i}^{\min(i+K-1,n)} y_{i,t}, \quad i = 1, \dots, n \tag{23}$$

$$\rho_i, m_i, sm_i, z_i \in \{0,1\}, \quad i = 1, \dots, n \tag{24}$$

$$y_{i,t} \in \{0,1\}, \quad i = 1, \dots, n,\ t = i, \dots, \min(i + K - 1, n) \tag{25}$$

$$S_i, C_i, C_i^{eff}, T_i, H_i \ge 0, \quad i = 1, \dots, n \tag{26}$$

Equation (1) is the objective function, which minimizes weighted job tardiness, maintenance, sampling, rejection, and non-validation costs. Constraints (2)–(3) establish the job chronology, while constraints (4) ensure that non-rejected jobs start no earlier than their release dates. Constraints (5) set the potential completion time for each job. Constraints (6)–(8) link the effective completion time (C_i^{eff}) to the actual completion time (C_i) for non-rejected jobs. This prevents the objective function from being erroneously minimized by rejecting jobs. Constraints (9) define the tardiness of each job.

Constraints (10)–(12) describe the evolution of the EHI, where the system health degrades linearly with processing time unless maintenance is performed. Constraints (13)–(14)

enforce that $H_i = EHI_{max}$ when maintenance is carried out. Constraints (15) ensure that a job is processed only if the EHI remains above the critical threshold θ. Constraints (16) prevent the insertion of maintenance immediately before a rejected job, since performing maintenance before a job that is not processed is operationally meaningless. Similarly, constraints (17) ensure that no retroactive sampling is activated after a rejected job.

Constraints (18)–(20) define the retroactive sampling coverage mechanism. Constraints (18) link the coverage variable $y_{i,t}$ to the activation of sampling at position t: a job i can be validated at t only if $sm_t = 1$. Constraints (19) prevent rejected jobs from being validated, as rejected lots do not require quality confirmation. Constraints (20) enforce the retroactive nature of the sampling window: sampling at position t may validate a preceding job i only if no maintenance occurs between positions i and t.

Constraints (21)–(23) define the global validation indicator. Constraints (21) enforce that a rejected job is always considered validated from a quality-risk perspective. Constraints (22) ensure that if a job is validated by at least one sampling, then $z_i = 1$. Constraints (23) impose that z_i can be equal to 1 only if the job is either rejected or validated by at least one sampling within the backward window.

Finally, constraints (24)–(25) specify the binary nature of the decision variables, while constraint (26) enforces nonnegativity of all continuous variables.

III. RESOLUTION APPROACH

This section presents the computational complexity analysis of the *FS-SM-K* problem. It is first shown that the problem is NP-complete, which justifies the use of the $MILP$ formulation for small instances and motivates the development of efficient solution approaches for practical instances.

A. NP-Completeness Result

Let the decision version of the *FS-SM-K* problem be defined as follows: given an instance of the problem and a bound B, does there exist a feasible solution with an objective value of at most B?

Theorem 1: The decision version of the *FS-SM-K* problem is NP-complete.

Proof: To establish this result, NP-hardness is proven by reduction from the decision version of the 0-1 KNAPSACK problem, and it is then shown that the decision version of *FS-SM-K* belongs to NP.

1) Sketch of NP-hardness: The NP-hardness is shown via a polynomial-time reduction from the 0-1 KNAPSACK problem, which is known to be NP-complete (see [6]).

An instance of the decision version of the 0-1 KNAPSACK problem is defined by a set of N items, each with a weight $a_i \in \mathbb{Z}^+$ and a value $v_i \in \mathbb{Z}^+$, a maximum weight capacity $W \in \mathbb{Z}^+$, and a target value $P \in \mathbb{Z}^+$. The question is whether there exists a subset of items to include in the knapsack such that the total weight does not exceed W, and the total value is at least P.

Given such an instance, a corresponding instance of the decision version of the *FS-SM-K* problem is constructed as follows:

- The number of jobs is $n = N$, and the fixed processing sequence is $\sigma = (1, \dots, n)$.
- For each job $i \in \{1, \dots, n\}$, the processing time is set to $p_i = a_i$, and the job weight in the objective function is set to $w_i = v_i$.
- The degradation rate is $\delta = 1$. The initial and maximum equipment health index is $EHI_{max} = W$, and the critical threshold is $\theta = 0$.
- Let $V = \sum_{i=1}^{N} v_i$ be the sum of all item values. The maintenance cost is set to a prohibitively large value $C_m > V$. The maintenance duration is $\tau_m = 0$.
- Retroactive sampling is rendered useless by setting the sampling cost $C_s = 0$ and the non-validation cost $C_{unv} = 0$. The sampling window size K can be arbitrarily set to 1. Since sampling does not affect feasibility and carries no cost, it can be ignored in any optimal solution.
- The rejection cost parameter is normalized to $C_r = 1$.
- Timing constraints are relaxed by setting all release dates $r_i = 0$ and all due dates $d_i = \infty$, so that the tardiness T_i is always 0.
- Finally, the target bound for the total cost is set to $B = V - P$.

The *FS-SM-K* decision question is whether there exists a feasible schedule with a total cost of at most B.

Under these settings, the objective function of the *FS-SM-K* problem, given by Equation (1), reduces to:

$$\min \sum_{i=1}^{n} w_i C_r \rho_i + C_m m_i = \min \sum_{i=1}^{N} v_i \rho_i + C_m m_i$$

where $\rho_i \in \{0, 1\}$ is the rejection decision for job i, and $m_i \in \{0, 1\}$ is the maintenance decision.

Since the target bound $B = V - P \leq V$ and $C_m > V$, any feasible solution with a cost of at most B must have $m_i = 0$ for all $i = 1, \dots, n$. In other words, no maintenance can be performed. Consequently, the objective function strictly measures the total value of the rejected jobs: $\sum_{i=1}^{N} v_i \rho_i$.

Since maintenance is excluded ($m_i = 0$) and $\delta = 1$, it follows by induction from constraints (10) and (12) that the equipment health index, immediately before processing job j, satisfies: $H_j = EHI_{max} - \sum_{i=1}^{j-1} p_i(1 - \rho_i)$.

For the schedule to be feasible, when a job j is accepted ($\rho_j = 0$), constraint (15) implies that $H_j \geq \theta$. Because $EHI_{max} = W$, $\theta = 0$, and $p_i = a_i$, this feasibility condition implies that the cumulative processing time of all accepted jobs across the entire sequence cannot exceed W: $\sum_{i=1}^{N} a_i(1 - \rho_i) \leq W$.

The equivalence between the two decision problems is now demonstrated.

Sense 1 (⇒): Suppose there exists a valid solution to the 0-1 KNAPSACK instance, i.e., a subset of items $S \subseteq \{1, \dots, N\}$ such that $\sum_{i \in S} a_i \leq W$ and $\sum_{i \in S} v_i \geq P$. A schedule for *FS-SM-K* is constructed by accepting job

i if item $i \in S$ (setting $\rho_i = 0$), and rejecting job i if item $i \notin S$ (setting $\rho_i = 1$). No maintenance is performed ($m_i = 0$). The total processing time of the accepted jobs is $\sum_{i \in S} a_i \leq W$, which ensures that the health index never drops below $\theta = 0$. The schedule is therefore feasible. The total cost of this schedule is the sum of the values of the rejected jobs: $\sum_{i \notin S} v_i = V - \sum_{i \in S} v_i$. Since $\sum_{i \in S} v_i \geq P$, the total cost is at most $V - P = B$. Thus, a *YES* instance of 0-1 Knapsack yields a *YES* instance of *FS-SM-K*.

Sense 2 ($\Leftarrow$): Conversely, suppose there exists a feasible solution for the constructed *FS-SM-K* instance with a total cost of at most B. As established earlier, this cost bound ensures that no maintenance is performed, and the cost strictly equals $\sum_{i=1}^{N} v_i \rho_i \leq B = V - P$. A solution for the Knapsack problem is constructed by selecting the items corresponding to the accepted jobs: $S = \{i \mid \rho_i = 0\}$. The feasibility of the schedule ensures that the total weight of these accepted jobs satisfies $\sum_{i \in S} a_i \leq W$. Furthermore, the total value of the selected items is $\sum_{i \in S} v_i = V - \sum_{i=1}^{N} v_i \rho_i$. Since $\sum_{i=1}^{N} v_i \rho_i \leq V - P$, it follows that $\sum_{i \in S} v_i \geq V - (V - P) = P$. Thus, a *YES* instance of *FS-SM-K* yields a *YES* instance of 0-1 Knapsack.

Note that the reduction only uses a restricted subclass of *FS-SM-K* instances (with no maintenance and no sampling). Therefore, the general *FS-SM-K* problem is at least as hard, which proves NP-hardness.

The above construction can be carried out in time linear in N. Therefore, this defines a polynomial-time reduction from 0-1 KNAPSACK to *FS-SM-K*, proving NP-hardness.

2) NP-membership:

Lemma 1: The decision version of *FS-SM-K* belongs to NP.

Given a candidate solution consisting of the binary decisions $(\rho_i, m_i, sm_i, y_{i,t}, z_i)$ and the corresponding schedule. Feasibility can be verified in polynomial time by checking all linear constraints of the $MILP$ formulation, including the timing constraints, the equipment health evolution and threshold conditions, and the retroactive validation coverage rules. Furthermore, the objective value can be computed in $O(n)$ time, and it can be verified in $O(1)$ time whether this value is at most B. Therefore, the decision version of *FS-SM-K* belongs to NP.

Since the decision version of *FS-SM-K* is NP-hard and belongs to NP, it follows that it is NP-complete.

■

IV. Experimental Study

This section presents the numerical experiments conducted to evaluate the performance of the proposed model, which is implemented using the *gurobipy* library and solved with GUROBI. All experiments are carried out on a computer equipped with an Intel(R) Core(TM) i5-1135G7 processor running at 1.1 GHz and 16 GB of RAM, under Windows 11. A CPU time limit of 3600 seconds is imposed. Before discussing the obtained numerical results, the method for generating instances is described below.

A. Benchmark description

The performance of P_1 is evaluated using randomly generated semiconductor manufacturing instances. Three instance sizes are considered: (i) small instances, with $n \in \{10, 20, 30\}$, (ii) medium instances, with $n \in \{50, 60, 70\}$, and (iii) large instances, with $n \in \{100, 110, 120\}$.

For the job-related parameters, processing times p_i, release dates r_i, and due dates d_i are uniformly generated from $[10, 30]$, $[0, 0.01 \times \sum_{j=1}^{n} p_j]$, and $[r_i + \sum_{j=1}^{i} p_j, r_i + \sum_{j=1}^{n} p_j]$, respectively. The job weights w_i follow a Pareto distribution, reflecting a heterogeneous priority structure in which a small proportion of jobs accounts for a large share of the total weight (approximately 80% of the weight concentrated on 20% of the jobs). This setting captures typical conditions in semiconductor manufacturing, where jobs within the fabrication facility have varying levels of priority, modeled here through their associated weights.

Regarding the maintenance parameters, $EHI_{max} = 100$, $\delta = 0.5$, and $\theta = 70$. Maintenance cost C_m, maintenance duration τ_m, and sampling C_s follow the uniform distribution from $[100, 500]$, $[10, 50]$, and $[100, 200]$, respectively. The non-validation cost is set to $C_{unv} = 2C_s$.

Furthermore, values of the rejection cost C_r are uniformly generated across three levels: (i) low, with $C_r \in [200, 400]$, (ii) medium, with $C_r \in [400, 700]$, and (iii) high, with $C_r \in [1200, 1500]$. Similarly, three levels of the parameter K are considered to distinguish between short, intermediate, and large retroactive validation windows: (i) short, with $K = 0.1n$, (ii) intermediate, with $K = 0.25n$, and (iii) large, with $K = 0.5n$. Overall, considering 9 values of n, 3 levels of C_r, and 3 levels of K, a total of 81 instances are generated.

B. Numerical results

The first part of the numerical experiments aims to evaluate the performance of P_1 and to gain insights into its behavior as the instance size increases. To this end, the computational time of P_1, expressed in seconds, is recorded for all instance. Then, for each pair (n, K), the average computational time is calculated over three corresponding instances (low C_r, medium C_r and high C_r). The obtained results are summarized in Figure 1 where each line corresponds to a specific value of K.

Results from Figure 1 show that the computational time of P_1 remains within seconds for small instances. A similar observation applies to medium instances, which are solved optimally within a few minutes. However, for large instances, the computational time of P_1 increases significantly for all values of K, reaching on average 2518 seconds, 3221 seconds, and 3240 seconds for the pairs ($n = 120$, short K), ($n = 120$, intermediate K), and ($n = 120$, large K), respectively. Notably, for half of the large instances, P_1 reaches the 3600-second time limit, indicating that the optimal solution was not found within the allowed time. This increasing in computational times highlights the intractability of large instances and motivates the development of heuristic methods to efficiently solve the addressed problem.

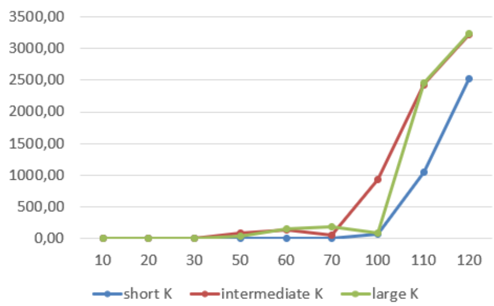

Fig. 1: Processing time of P_1 in seconds.

The second part of the numerical experiments studies the sensitivity of the model P_1 to variations in key parameters. In particular, it focuses on how different scheduling criteria affect job rejection. To this end, P_1 is modified to minimize completion time instead of total tardiness. This involves replacing w_iT_i in the objective function with $w_iC_i^{eff}$, and removing the variable T_i along with constraints (9). For all instances, the rejection rate is computed under both scheduling objectives, w_iT_i and $w_iC_i^{eff}$, as the ratio $\frac{\text{Total number of the rejected jobs}}{n}$. The average rejection rate is then computed for each pair (n, C_r) across the three levels of sampling scope K (short, intermediate, and large). Figure 2 shows the obtained results for different levels of C_r.

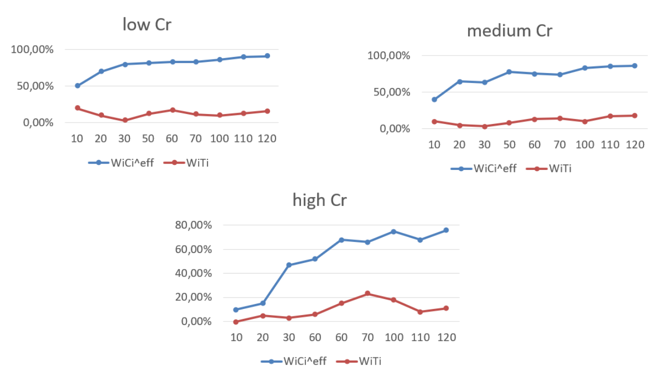

Fig. 2: Rejection rate: w_iT_i and $w_iC_i^{eff}$.

An analysis of Figure 2 reveals that the rejection rate is consistently higher under the $w_iC_i^{eff}$ objective than under w_iT_i for all instances. This can be explained by the fact that $w_iC_i^{eff}$ increases rapidly as the job sequence progresses, which encourages the model to reject more jobs in order to limit the growth of the objective function. This behavior highlights the strong influence of the scheduling criterion on rejection decisions and underlines the trade-off between minimizing completion times and controlling the number of rejected jobs.

V. CONCLUSION AND PERSPECTIVES

This paper addresses an integrated scheduling problem on a single machine within a semiconductor manufacturing environment, where equipment degradation is explicitly modeled alongside product quality considerations. Given a fixed job sequence, the problem involves determining optimal decisions regarding preventive maintenance, retroactive quality sampling, and job rejection, with the objective of minimizing total tardiness and operational costs. The problem is formulated as a $MILP$ model that incorporates constraints related to sequencing, equipment degradation, maintenance, and retroactive quality sampling. Furthermore, the NP-hardness of the problem is established through theoretical analysis. Computational experiments, based on randomly generated instances reflecting semiconductor manufacturing settings, showed that the proposed model becomes intractable for large-scale instances. The results also highlighted the significant influence of scheduling-related criteria on the model behavior, particularly with respect to job rejection rates.

This study opens several directions for future research. First, given the NP-hard nature of the problem, the development of efficient heuristics and metaheuristics represents a promising direction. Second, a potential extension would be to relax the assumption of a fixed job sequence, allowing scheduling decisions to be integrated directly into the model. Finally, extending the model to incorporate non-linear degradation dynamics could lead to a more realistic representation and additional insights into the model results.

REFERENCES

[1] E. H. Aghezzaf, M. A. Jamali, and D. Ait-Kadi, "An integrated production and preventive maintenance planning model," *European journal of operational research*, vol. 181, no. 2, pp. 679–685, 2007.

[2] R. Ahmad and S. Kamaruddin, "An overview of time-based and condition-based maintenance in industrial application," *Computers & Industrial Engineering*, vol. 63, no. 1, pp. 135–149, August 2012.

[3] C. R. Cassady and E. Kutanoglu, "Integrating preventive maintenance planning and production scheduling for a single machine," *IEEE Transactions on reliability*, vol. 54, no. 2, pp. 304–309, 2005.

[4] G. Q. Cheng, B. H. Zhou, and L. Li, "Integrated production, quality control and condition-based maintenance for imperfect production systems," *Reliability Engineering & System Safety*, vol. 175, pp. 251–264, July 2018.

[5] C.-F. Chien, T. H. V. Nguyen, Y.-C. Li, and Y.-J. Chen, "Bayesian decision analysis for optimizing in-line metrology and defect inspection strategy for sustainable semiconductor manufacturing and an empirical study," *Computers & Industrial Engineering*, vol. 182, p. 109421, August 2023.

[6] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman, 1979.

[7] C.-Y. Lee and Z.-H. Dong, "Hierarchical equipment health index framework," *IEEE Transactions on Semiconductor Manufacturing*, vol. 32, no. 3, pp. 267–276, August 2019.

[8] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," *Manufacturing letters*, vol. 3, pp. 18–23, 2015.

[9] G. S. May and C. J. Spanos, *Fundamentals of semiconductor manufacturing and process control*. John Wiley & Sons, 2006.

[10] L. Mönch, J. W. Fowler, S. Dauzère-Pérès, S. J. Mason, and O. Rose, "A survey of problems, solution techniques, and future challenges in scheduling semiconductor manufacturing operations," *Journal of scheduling*, vol. 14, no. 6, pp. 583–599, 2011.

[11] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation–a review on the statistical data driven approaches," *European journal of operational research*, vol. 213, no. 1, pp. 1–14, 2011.

[12] C.-W. Wu and Z.-H. Wang, "A cost-effective sampling strategy by variables inspection for lot disposition," *Annals of Operations Research*, vol. 349, pp. 219–235, 2025, published online 2023.

[13] C. Yugma, J. Blue, S. Dauzère-Pérès, and A. Obeid, "Integration of scheduling and advanced process control in semiconductor manufacturing: review and outlook," *Journal of Scheduling*, vol. 18, no. 2, pp. 195–205, 2015.