



Gong, X., Liu, Y. , Lohse, N., De Pessemier, T., Martens, L. and Joseph, W. (2018) Energy and labor aware production scheduling for industrial demand response using adaptive multi-objective memetic algorithm. *IEEE Transactions on Industrial Informatics*,(doi:[10.1109/TII.2018.2839645](https://doi.org/10.1109/TII.2018.2839645))

This is the author's final accepted version.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/163535/>

Deposited on: 08 June 2018

Enlighten – Research publications by members of the University of Glasgow  
<http://eprints.gla.ac.uk>

# Energy- and Labor-Aware Production Scheduling for Industrial Demand Response Using Adaptive Multiobjective Memetic Algorithm

Xu Gong, Ying Liu, Niels Lohse, *Member, IEEE*, Toon De Pessemer, Luc Martens, *Member, IEEE*, and Wout Joseph, *Senior Member, IEEE*

**Abstract**—Price-based demand response stimulates factories to adapt their power consumption patterns to time-sensitive electricity prices, so that a rise in energy cost is prevented without affecting production on the shop floor. This paper introduces a multiobjective optimization (MOO) model that jointly schedules job processing, machine idle modes, and human workers under real-time electricity pricing. Beyond existing models, labor is considered due to a common trade-off between energy cost and labor cost. An adaptive multiobjective memetic algorithm (AMOMA) is proposed to fast converge toward the Pareto front without loss in diversity. It leverages feedback of cross-dominance and stagnation in a search and a prioritized grouping strategy. In this way, adaptive balance remains between exploration of the nondominated sorting genetic algorithm II (NSGA-II) and exploitation of two mutually complementary local search operators. A case study of an extrusion blow molding process in a plastic bottle manufacturer and benchmarks demonstrate the MOO effectiveness and efficiency of AMOMA. The impacts of production-prohibited periods and relative portion of energy and labor costs on MOO are further analyzed, respectively. This method was further generalized in a multi-machine experiment. The common trade-off relations between the energy and labor costs as well as between the makespan and the sum of the two cost parts were quantitatively revealed.

**Index Terms**—Demand-side management, evolutionary computation, intelligent manufacturing systems, multiobjective optimization, scheduling

## NOMENCLATURE

$C_{max}$	Makespan of the entire production
$CC_i$	Electricity cost for performing the $i$ -th changeover
$CI_i$	Electricity cost for performing the $i$ -th idling
$CJ_i$	Electricity cost for processing the $i$ -th job
$D$	Duration of the electricity pricing slot
$D_{poff}$	Duration of powering off a machine
$D_s$	Duration of machine power state $s$

X. Gong was with the Department of Information Technology, Ghent University, Belgium. He is now with the Department of Machine Intelligence Technology, DAMO Academy, Alibaba Group, Hangzhou, China (e-mail: xu.gong@outlook.com).

Ying Liu is with the School of Engineering, University of Glasgow, Glasgow, U. K. (e-mail: ying.liu@glasgow.ac.uk).

Niels Lohse is with the Wolfson School of Mechanical, Electrical and Manufacturing Engineering, Loughborough University, Loughborough, U. K. (e-mail: n.lohse@lboro.ac.uk).

T. De Pessemer, L. Martens, and W. Joseph are with the Department of Information Technology, Ghent University, Belgium (e-mail: toon.depessemier@ugent.be;luc1.martens@ugent.be; wout.joseph@ugent.be).

$DC_i$	Duration of the $i$ -th machine changeover
$DJ_j$	Duration of processing job with ID $j$
$DT$	Common due time of all jobs
$EP_{ts}$	Electricity price on $ts$ -th time slot
$ETC_i^n$	End time ( $\delta t$ ) of $n$ -th subpart of $i$ -th changeover
$ETJ_i^n$	End time ( $\delta t$ ) of $n$ -th subpart of $i$ -th job
$ETSC_i^n$	End time (pricing slots) of $n$ -th subpart of $i$ -th changeover
$ETSJ_i^n$	End time (pricing slots) of $n$ -th subpart of $i$ -th job
$N_J$	Number of jobs to be scheduled
$N_f$	Number of first generations without exploitation
$NSC_i$	Number of subparts of the $i$ -th changeover
$NSJ_i$	Number of subparts of the $i$ -th job
$P_p$	Power of the machine power state ‘Production’
$P_s^t$	Power of the machine power state $s$ at time $t$ (in $\delta t$ )
$pt$	Type of personnel required in shift $sh$
$S_c$	Sequence of power states for a machine changeover
$S_o$	Sequence of power states for switching to, staying at, and recovering from off between contiguous jobs
$S_{is}$	Tabu search step on the scheduling time horizon
$sh$	Labor shift corresponding to time $t$ (in $\delta t$ )
$SHIFT$	Labor shift types on a weekday or a weekend day
$SI_{ij}$	$SJ_j$ following the $i$ -th job
$SI_j$	Sequence of power states for switching to, staying at, and recovering from the $j$ -th machine idle mode
$STC_i^n$	Start time ( $\delta t$ ) of $n$ -th subpart of $i$ -th changeover
$STJ_i^n$	Start time ( $\delta t$ ) of $n$ -th subpart of $i$ -th job
$STSC_i^n$	Start time (pricing slots) of $n$ -th subpart of $i$ -th changeover
$STSJ_i^n$	Start time (pricing slots) of $n$ -th subpart of $i$ -th job
$t$	Absolute time or clock time
$TC$	Total cost of the production load
$TEC$	Total energy cost of the production load
$TLC$	Total labor cost of the production load
$T_{max}$	Maximal number of stagnated generations
$T_r$	Common release time of all jobs
$W_{sh}^{pt}$	Labor wage of the $pt$ in the $sh$
$\delta sh$	Duration of one labor shift
$\delta t$	Scheduling time slot
$\lambda$	Binary production-prohibited period indicator
$\theta_{sh}^{pt}$	Boolean indicator for the $pt$ in the $sh$
$\beta_{ts}$	Binary time slot indicator
$\pi$	Job sequence

## I. INTRODUCTION

Price- or time-based demand response (DR) stimulates end users to adapt their electricity consumption patterns to time-sensitive electricity prices. In this way, end users could possibly reduce their energy cost and the grid stability can be enhanced by a better-balanced demand-supply [1]. The price-based DR has focused on the residential domain, such as scheduling electrical loads of household appliances [2] together with electric vehicles under real-time electricity pricing [3]. It recently penetrates the manufacturing industry, which is a major electricity consumer. The industrial price-based DR can be realized by energy-aware scheduling of production processes [4-6]. These methods perform production load/job shifting under real-time electricity pricing, while setting machines to off or standby modes during inactive periods.

In the residential DR [2, 3], household appliances, e.g., a dishwasher and a washing machine, can automatically perform the scheduled operations without human interventions. Comparatively, the human worker remains an indispensable factor on the shop floor, though robots are introduced to automate production to different degrees in factories [7]. It is then of practical significance to jointly schedule human workers and machine operations when factories perform the industrial DR for economical production. Unfortunately, the existing methods widely ignore the interdependence of energy and labor costs [8]. Shifting production loads from a day to a night or a weekend for energy cost reduction is usually compromised by an increased labor wage and thus a rising labor cost [9-11]. Consequently, the overall production cost may risk to rise.

According to the authors' cost breakdown of a plastic bottle manufacturer, which has a number of independent extrusion blow molding (EBM) machines, the labor cost is over 3 times higher than the energy cost. As a result, the former is much more sensitive to load shifting than the latter. Though the portion of two cost parts differs on a case-by-case basis, it is of economic benefits to jointly consider both energy and labor costs in production scheduling for the industrial DR, and study it as a multiobjective optimization problem (MOP), instead of only integrating energy awareness in conventional production scheduling algorithms in most existing studies.

Both electricity and labor costs were considered in the optimization of a multi-pass face milling process [11]. However, both costs were calculated using flat rates. These two cost parts were explicitly modeled in the flow shop scheduling problem under time-varying electricity and labor pricing [9]. Nevertheless, the following limitations are observed: (1) the labor rate does not vary among workers; (2) the production-prohibited period, which is often introduced by the labor shift, and its constrained influence on production operations are not modeled; (3) the exhaustive search is a rude solution method with poor scalability. A similar problem was investigated in [10] in a single objective optimization manner. Consequently, the trade-off relations in these cost parts and other important production metrics were not quantified, besides the aforementioned ignorance of production-prohibited periods.

Multiobjective evolutionary algorithms (MOEAs) are practical to solve multiobjective production scheduling

problems [12], as they are characterized by finding high-quality solutions in reasonable time, fitting the requirement of production scheduling [13]. Scalarization-based MOEAs [14] transform a MOP to a single objective optimization problem by summing weighted objectives in one fitness function. Nonetheless, it is problematic to assign proper weights for Pareto front (PF) approximation. Comparatively, domination-based MOEAs have been widely proven to be effective, among which NSGA-II [15] is highly representative.

Beyond using a decent NSGA-II to produce nondominated solutions [16], recent studies hybrid with one or more local searches to accelerate the convergence rate toward the PF without loss in diversity [17]. While a genetic search explores the solution space for potential regions, a local search exploits these regions by incorporating domain-specific knowledge on how a solution can be further improved. Such a hybrid is named as a memetic algorithm (MA) [18]. In simple MAs, domain knowledge is only captured and incorporated once by human experts at the design phase. Adaptive MAs additionally integrate knowledge on how an instance of MA (MA on the fly) is self-reconfigurable to better suit the problem when a search progresses [19]. Compared to employing MAs for many unconstrained optimization problems, few studies tackled multiobjective constrained problems by adaptive MAs [19, 20].

This paper proposes an adaptive multiobjective MA (AMOMA) based on the NSGA-II to optimize an integrated energy- and labor-aware production scheduling model under real-time electricity pricing. The contributions of this paper are threefold. (1) Compared to existing energy-aware production scheduling models, the proposed model additionally considers the labor type and quantity, the work shift, as well as the production-prohibited periods. This makes industrial load shifting more realistic. (2) The proposed AMOMA synergistically integrates in the NSGA-II convergence- and diversity-oriented tabu searches (TSs), respectively. It further adaptively coordinates the exploration and the exploitation during a search. (3) A case study on an EBM machine is performed. Using empirical data and extensive benchmarks, the proposed AMOMA are proven to achieve fast PF approximation while preserving diversity for this highly-constrained MOP.

## II. ENERGY- AND LABOR-AWARE SCHEDULING MODEL

The problem is to perform cost-effective production scheduling before a due time ( $DT$ ) and under real-time pricing (RTP) as well as the time-varying labor wage, aiming to minimize the following bi-objectives:

$$\min_{\pi, STJ_i^n, s, pt, sh} (TC, C_{max}), \text{ where } TC = TEC + TLC \quad (1)$$

where the total cost ( $TC$ ) comprises the total energy cost ( $TEC$ ) and the total labor cost ( $TLC$ ). While the existing relevant models focus on enabling energy awareness in scheduling, energy- and labor-related decision variables are integrated in this model: (1) the job sequence ( $\pi$ ), (2) the job start time ( $STJ_i^n$ ) considering the production-prohibited periods which may split a job into multiple subparts, (3) the machine power states ( $s$ ), (4) the number of each type of personnel ( $pt$ ), and (5)

the labor shift ( $sh$ ).  $\pi$  and  $STJ_i^n$  define job sequencing and timing, respectively;  $s$  assigns machine power states for job processing and idling (off and standby) between jobs; both  $pt$  and  $sh$  are adaptively determined according to the scheduled production. This integrated decision making is crucial to reduce the overall production cost, without neglecting the dependency between the  $TEC$  and the other important production metrics.

The production cost often consists of the machine depreciation cost, the material cost, the energy cost, the labor cost. It should be minimized in order to maximize the profit of a manufacturer. The number, type, release time, and due time of production jobs, which are predetermined by production planning [21], are the input variable for this scheduling problem. Therefore, the machine depreciation is considered as a fixed cost on the fixed short scheduling horizon with the fixed amount of production. The material cost increases linearly with the amount of production and cannot be influenced by the manufacturer [9]. Comparatively, the energy and labor costs are the key production cost parts whose variance is directly linked to the production. As a result, this energy- and labor-aware production scheduling problem enables profit maximization.

Compared to the prevalent residential DR studies [2, 3] that basically determine the simple operations (on/off) of household appliances in each scheduling time slot, this industrial DR model has threefold contributions. (1) Besides the decision-making of multiple machine operations (processing, on, off, and multiple idle modes), it integrates job sequencing and timing as well as human worker and labor shift planning. (2) The time granularity is reduced (second-scale or even smaller) for finer-grained scheduling and analytics. (3) A complete state-based energy model is employed for a more realistic consideration of the energy consumption behavior of a machine.

#### A. Total Labor Cost (TLC)

The  $TLC$  depends on  $sh$  and calculated by Eq. (2), where a labor wage  $W_{sh}^{pt}$  is determined by the types of both personnel and shift. If a  $pt$  is required by an involved power state, it is included in the corresponding shift (i.e.,  $\theta_{sh}^{pt} = 1$ ). Otherwise,  $\theta_{sh}^{pt} = 0$ . Therefore, a human worker is paid based on the occupied shifts. While a scheduling time span is divided into a number of shifts, the  $TLC$  is accumulated over these shifts and the human workers arranged in each shift.

$$TLC = \sum_{sh=ST}^{DT} \sum_{pt \in PT} (W_{sh}^{pt} \theta_{sh}^{pt}) \quad (2)$$

The duration of a shift ( $\delta sh$ , in hours) is defined in Eq. (3). While  $\delta sh$  remains constant, the shift types in a day ( $SHIFT$ ) depend on the time in a day as well as weekdays and weekends.

$$\delta sh = 24 / |SHIFT| \quad (3)$$

The purpose of this labor model is to plan the type and the number of human workers in each shifts, so as to enable the total labor cost-involved multiobjective optimization. Therefore, it is assumed that there are sufficient human workers. Specific labor factors, which can be further integrated to this labor model, are not considered, e.g., consecutive shifts for the same worker.

#### B. Total Energy Cost (TEC)

Under RTP, the electricity price varies across pricing time slots, but stays constant within each slot ( $D$ ). Consequently, the  $TEC$  varies with load shifting. It comprises the energy cost for processing jobs, as well as performing machine changeovers and idling between jobs, as formulated in Eq. (4).

$$TEC = \sum_{i=1}^{N_j} CJ_i + \sum_{i=1}^{N_j-1} (CC_i + CI_i) \quad (4)$$

As described by Eq. (5), the energy cost for the  $i$ -th job ( $CJ_i$ ) includes that for processing each subpart of this job if one or more production-prohibited periods split this job into multiple subparts ( $\lambda = 1$ ), and that for the  $NSC_i-1$  machine power-off&on operations if  $\lambda = 1$ . A job has only one subpart if  $\lambda = 0$ . This calculation method analogously applies to the energy cost for the  $i$ -th changeover ( $CC_i$ ) in Eq. (6).

$$CJ_i = \sum_{n=1}^{NSJ_i} \sum_{ts=STJ_i^n}^{ETJ_i^n} EP_{ts} \left( \sum_{t=STJ_i^n}^{ETJ_i^n} \beta_{ts} P_p^t \right) + \lambda \sum_{n=1}^{NSJ_i-1} \sum_{ts=STJ_i^n}^{ETJ_i^n} EP_{ts} \left( \sum_{t=ETJ_i^n}^{STJ_i^{n+1}} \sum_{s \in S_o} \beta_{ts} P_s^t \right), \quad (5)$$

$$i \in [1, 2, \dots, N_j]$$

$$CC_i = \sum_{n=1}^{NSC_i} \sum_{ts=STSC_i^n}^{ETSC_i^n} EP_{ts} \left( \sum_{t=STSC_i^n}^{ETSC_i^n} \sum_{s \in S_c} \beta_{ts} P_s^t \right) + \lambda \sum_{n=1}^{NSC_i-1} \sum_{ts=STSC_i^n}^{ETSC_i^n} EP_{ts} \left( \sum_{t=ETSC_i^n}^{STSC_i^{n+1}} \sum_{s \in S_o} \beta_{ts} P_s^t \right), \quad (6)$$

$$i \in [1, 2, \dots, N_j - 1]$$

The energy cost for machine idling encompasses the period when a machine switches to and stays at one of multiple idle modes, and returns to production state, as defined in Eq. (7). More specifically, the  $i$ -th machine idling starts from the end of the last subpart of the  $i$ -th job ( $ETJ_i^{NSJ_i}$  or  $ETJ_i^{NSJ_i}$ ) and ends at the start of the first subpart of the following  $i$ -th changeover ( $STC_i^1$  or  $STSC_i^1$ ). The absolute time is mapped to the electricity pricing slot by Eqs. (8-9).

$$CI_i = \sum_{ts=STSC_i^1}^{STSC_i^1} EP_{ts} \left( \sum_{t=ETJ_i^{NSJ_i}}^{STC_i^1} \sum_{s \in S_{ij}} \beta_{ts} P_s^t \right), \quad (7)$$

$$i \in [1, 2, \dots, N_j - 1]$$

$$\beta_{ts} = \begin{cases} 1, & \text{if } t \in [ts \cdot D, (ts+1) \cdot D) \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$ts = \lfloor (t - T_r) / D \rfloor, t \in [T_r, T_r + \delta t, \dots, DT - \delta t, DT] \quad (9)$$

#### C. Job and Changeover

Jobs can follow an arbitrary sequence with a common release time  $T_r$  and  $DT$ . Jobs may have different release and due times in more general scenarios. Nevertheless, this assumption is sufficient to enable job sequencing and timing, which are two essential decision variables even in these general scenarios. Due to the weekend production constraint introduced by the labor model, a job contains one or more sub-durations, as indicated in Eq. (10).

$$DJ_j = \sum_{n=1}^{NSJ_j} (ETJ_j^n - STJ_j^n), \quad i \in [1, 2, \dots, N_j] \quad (10)$$

The last subpart of the last job must be completed before  $DT$ , considering the duration to power off a machine (Eq. 11).

$$ETJ_{N_j}^{NSJ_{N_j}} + D_{poff} \leq DT, j \in [1, 2, \dots, N_j] \quad (11)$$

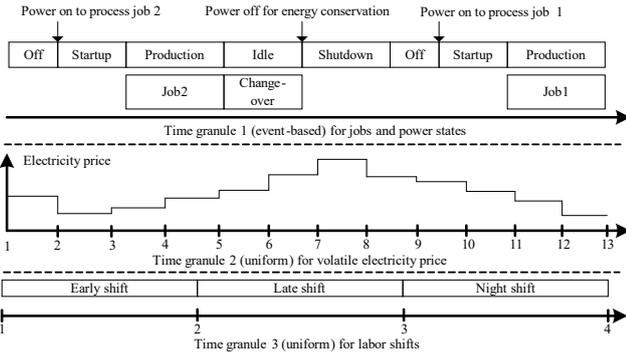


Fig. 1. Parallel time axes to synchronize electricity prices, jobs, shifts, power states, and machine operations in discrete-event simulation of a schedule.

As defined by Eq. (12), a changeover is required between adjacent jobs and starts right before the upcoming  $(i+1)$ -th job, i.e., the last subpart of the  $i$ -th changeover ends at the start time of the first subpart of the  $(i+1)$ -th job.

$$ETC_i^{NSC_i} = STJ_{(i+1)}^1, i \in [1, 2, \dots, N_J - 1] \quad (12)$$

Analogously, the duration of a changeover is the sum of potentially multiple sub-durations, as described in Eq. (13).

$$DC_i = \sum_{n=1}^{NSC_i} (ETC_i^n - STC_i^n), i \in [1, 2, \dots, N_J - 1] \quad (13)$$

#### D. Machine

A machine is assumed to have sufficient material supply and have no breakdown in order to focus on static scheduling. For practical dynamic scenarios, e.g., shortage of material supply and machine failure during the execution of a production schedule, rescheduling is a common method [22]. However, the static scheduling techniques can be leveraged to enable the two popular rescheduling methods, i.e., full generation and repair of a schedule [22].

A machine cannot simultaneously process multiple jobs and does not allow preemption (Eq. 14). As formulated in Eq. 15, an idle mode is only applicable to an inter-job period that can accommodate it. Such an inter-job period is between the end of the last subpart of the  $i$ -th job ( $ETJ_i^{NSJ_i}$ ) and the start of the first subpart of the  $i$ -th changeover ( $STC_i^1$ ). If the production is prohibited during a period, the machine must stay off during this period, as indicated in Eq. (16).

$$ETJ_i^{NSJ_i} \leq STJ_{(i+1)}^1, i \in [1, 2, \dots, N_J - 1] \quad (14)$$

$$\sum_{s \in SI_j} D_s \leq STC_i^1 - ETJ_i^{NSJ_i}, i \in [1, 2, \dots, N_J - 1] \quad (15)$$

$$P_s^t = 0, \text{ if } (\lambda = 1) \ \& \ (\forall t \in \text{production-prohibited periods}) \quad (16)$$

### III. ADAPTIVE MULTIOBJECTIVE MEMETIC ALGORITHM

Considering the time-critical application of production scheduling under real-time pricing, a fixed time budget is given to the AMOMA, of which a practical value is 3 min [12]. This is fast in the sense that both RTP and labor shifts have hour-based granularities, such that machines and human workers may roughly have hour-based time to adjust to a new schedule which is adapted to the time-varying electricity prices and labor wage. However, this is challenging given the highly-

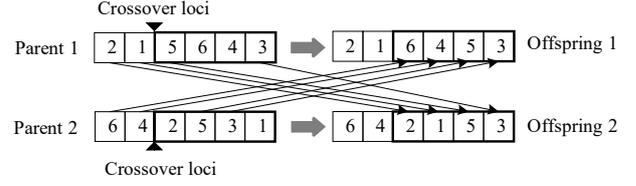


Fig. 2. One-point crossover where the number represents the job ID.

constrained MOP, a potentially huge solution space, and a preferred ordinary computer due to the cost issue.

#### A. Solution Encoding and Fitness Evaluation

A schedule is encoded as a chromosome which has a sequence of jobs as genes. A gene is identified by job ID and contains decision variables for this job: start time, changeover before this job, and machine idle mode following this job.

Regardless of the fitness assignment strategy, an essential work of fitness evaluation is to fast calculate  $TEC$ ,  $TLC$ , and  $C_{max}$  of a solution, so more computation resources can be allocated to the search process. As finite state machines (automata) [6] enable discrete-event simulation, they fit this requirement. Once power data is collected from a machine, energy profiling is performed to get a complete set of power states with the corresponding power profile (averaged power and duration).  $TEC$  and  $TLC$  are accumulated with state retention and transition over time, while  $C_{max}$  equals the time when the entire production is completed. The triggering events for power state transition include “power on/off the machine”, “start production”, and “select an idle mode”, whose timestamps are assigned by a schedule.

A time-granule-based method is elaborated to synchronize various elements in discrete-event simulation of a schedule. As illustrated in Fig. 1, three parallel time axes with independent time granules are coupled. Time axis 1 synchronizes jobs and power states, where the time elapse is triggered by machine operations. Time axis 2 coordinates real-time electricity prices that can be obtained from the electricity spot market beforehand, with the time granule of pricing period. Time axis 3 matches labor shifts. Its time granule is the duration of a shift.

#### B. Exploration by Genetic Search

The exploration framework employs the prevalent genetic search NSGA-II [15], which uses domination as the fitness assignment strategy. It searches for potential regions in the solution space without having to guarantee local optimum in each region. Thereby, exploration equals diversity preservation in AMOMA. As two local search operators are used (Sect. III-C2), the population size should stay sufficiently large (larger than that for pure exploration) to balance the computation resource for exploration and exploitation in a generation.

Several measures are taken on the crossover to preserve the diversity. First, parents are chosen via a binary tournament selection, preserving the diversity to the maximal extent. Second, a one-point crossover (Fig. 2) is employed, such that a solution can change the job sequence ( $\pi$ ) combing partial sequences of parents. As  $TEC$ ,  $TLC$ , and  $C_{max}$  are sensitive to job shifting, crossover loci are randomly selected and offspring are randomly timed (without altering  $\pi$ ). Third, the crossover rate remains high to introduce sufficient recombination of

TABLE I

ANALYSIS OF CONVERGENCE- AND DIVERSITY-ORIENTED TABU SEARCH (CTS AND DTS) IN OPTIMIZING TOTAL COST ( $TC$ ) AND MAKESPAN ( $C_{max}$ )

Tabu search	$TC$	$C_{max}$	Convergence	Diversity	Complexity <sup>b</sup>
CTS	↓	- <sup>a</sup> or ↓	↑	↓	$O(mn^2)$
DTS	↑	↓	-	↑	$O(mn)$

<sup>a</sup> -: no impact. <sup>b</sup> $m$ : number of time slots,  $n$ : number of jobs

solutions and thus a higher opportunity for the genetic search to enter diverse areas.

A swap mutation is used for a solution to switch two of its randomly selected jobs. The diversity is preserved by two measures. First, random timing is performed on all jobs after a mutation (without altering  $\pi$ ). Second, as a mutation follows a crossover, the mutation rate remains low, to maintain the effect of crossover and to avoid a pure random search.

Fully feasible solutions are produced to remove the need for repairing infeasible solutions. This is realized by assigning the start time of one job after another, according to the natural order of scheduled job positions. When timing a job, a random start time is generated from its maximal slack regarding the  $DT$ . Furthermore, redundancy is fully prevented to increase the diversity and also to reduce a waste of computation resources. A solution is considered redundant if it equals another in the multiobjective space. Reproduction is iteratively performed upon redundancy until it is removed.

### C. Exploitation by Multiple Memes

Mememes are incorporated in the exploitation framework of AMOMA in two manners: preprocessing scheme and problem-specific local search operators. The former integrates a priori knowledge in population initialization, biasing the overall search from the start to promising regions. The latter is interwoven with genetic operators (Fig. 3) and leverages domain knowledge to refine selected solutions.

1) *Preprocessing*: Two dispatching rules, which match  $\{1, RTP\}$  split into  $\{TEC, TLC, C_{max}\}$ , introduce specialized solutions in initialization. (1) “As-early-as-possible”: all jobs are joint and start from the beginning such that  $C_{max}$  is minimized. (2) “As-late-as-possible”: all jobs are joint and start late such that the last job ends at the  $DT$  and  $C_{max}$  is maximized.

2) *Local Search Operators*: Two tabu search (TS) algorithms are proposed for local refinement: CTS and DTS. They are preferred over hill climbing, since they can escape a local optimum by temporarily accepting a deteriorated solution and potentially leading to a superior solution. They are mutually complementary by stimulating the convergence and diversity of nondominated solutions, respectively.

To enable exploitation of all neighborhoods, a greedy termination criterion is defined: the longest free period is traversed, among free periods that are inter-job, and before and after the entire production. The TS step ( $S_{ts}$ , basic time slot to define a neighborhood structure) thus determines the shared portion of genetic and local searches in a fixed time budget.

A neighborhood structure in a CTS is built through backward moving a block of  $n$  ( $n \geq 1$ ) contiguous jobs by one  $S_{ts}$ . In every TS iteration, blocks of jobs are constructed by starting from the first job and ending at every following job, restarting from the second job and ending at every following job, and so on until all possible neighbors are iterated. The best neighbor has the lowest  $TC$ , and the shortest  $C_{max}$  if multiple neighbors have an

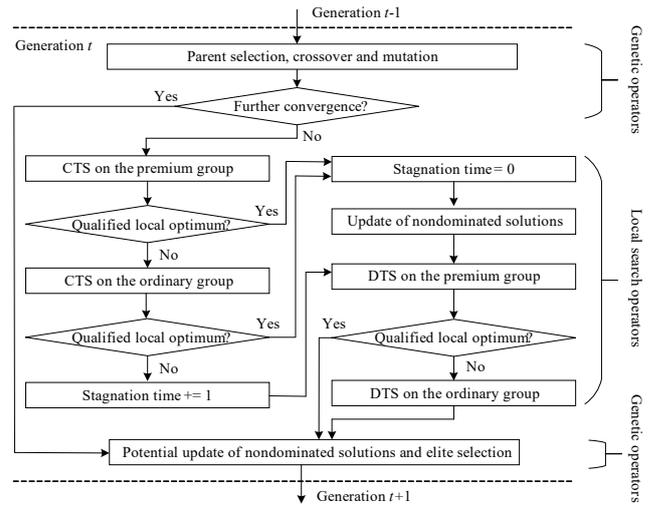


Fig. 3. Adaptive coordination of genetic search and two local search operators, i.e., convergence- and diversity-oriented tabu searches (CTS and DTS).

equal lowest  $TC$ . The aspiration criterion requires a neighbor to dominate at least one nondominated solution of the latest generation  $t$  ( $NS_t$ ). It is a soft criterion during the progress of a CTS instance, by only filtering neighbors when at least one qualified neighbor exists. It is a hard criterion at the end of a CTS instance, to ensure that the refined solution actually improves the convergence of nondominated solutions. As a CTS naturally reduces  $TC$  without increasing  $C_{max}$ , it enhances the convergence while losing the diversity (Table I) due to the concentration behavior of CTS.

In a DTS, a neighborhood structure is built via backward moving the last  $n$  ( $n \geq 1$ ) contiguous jobs by one  $S_{ts}$ . The last job must be included to reduce  $C_{max}$ . During each TS iteration, the best neighbor is the one that is not dominated by any other neighbors, and leads to the most evenly spread approximation set in case of multiple nondominated neighbors. The metric  $\Delta$  is used to indicate this evenness:

$$\Delta = \frac{1}{d} \sqrt{\frac{1}{|NS_t|} \sum_{x_i \in NS_t} (d_i - \bar{d})^2} \quad (17)$$

where  $d_i$  is the Euclidean distance between solution  $x_i$  and its nearest neighbor in  $NS_t$ , and  $\bar{d}$  is the mean Euclidean distance. A smaller  $\Delta$  implies a higher extent of spread in  $NS_t$ .  $\Delta$  has a complementary role of the crowding distance in NSGA-II for diversity preservation. The former smoothens the approximated PF, whereas the latter produces distant or extreme solutions by preferring less-crowded regions. The aspiration criterion requires a neighbor to be neither dominated by nor equal to any latest nondominated solutions. Analogously, it is a soft and hard criterion during and at the end of a DTS instance, respectively. As a DTS reduces  $C_{max}$  and increases  $TC$ , it enhances the diversity without influencing the convergence (Table I).

Fig. 3 depicts the synergy of genetic and local searches. The CTS is first applied after a genetic search. It contributes to the convergence of  $NS_t$  while deteriorating the diversity (Table I). The DTS then compensates by improving the diversity, in order to prevent the premature convergence (Table I). Both TSs are

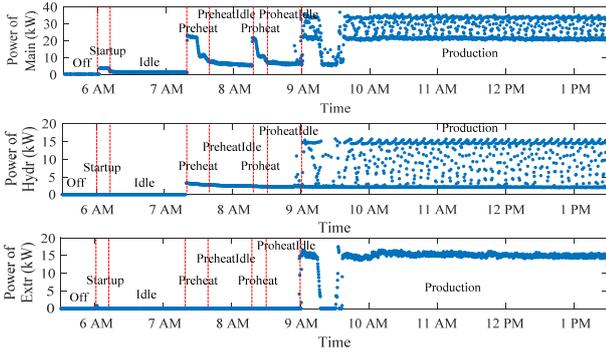


Fig. 4. Measured power data and power profile identification of three major energy consumers of an extrusion blow molding machine (from the top to the bottom: main system, hydraulic system, and extruder).

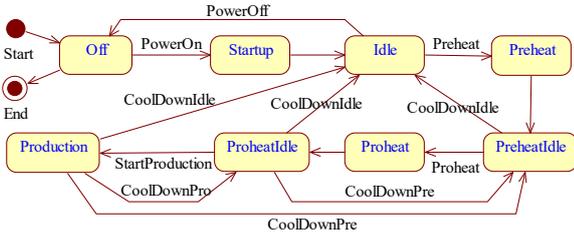


Fig. 5. Energy model of a continuous extrusion blow molding machine.

tabu list free. This is because they are mono-directional on the time span and intrinsically skip previously-visited solutions. A tabu list is thus no more needed to forbid the search direction.

D. Adaptive Coordination of Genetic and Local Searches

As preprocessing (Sect. III-C1) may lead to premature convergence upon the start of an entire search, local search operators are not utilized in the first  $N_f$  generations, where  $N_f$  controls the frequency of pure genetic search. Afterward, they are launched once  $NS_t$  that are output by the genetic search stops converging compared to  $NS_{t-1}$  (nondominated solutions in the previous generation  $t-1$ ). The cross-dominance metric  $\lambda$  [23] is employed to characterize this relative convergence:

$$\lambda = \Lambda_t / (|NS_t| \cdot |NS_{t-1}|) \tag{18}$$

where  $\Lambda_t$  denotes the number of dominance occurrences obtained by pairwise comparing  $NS_t$  to  $NS_{t-1}$ .  $\lambda$  of zero indicates that  $NS_t$  do not converge any more beyond  $NS_{t-1}$ .

Two groups of solutions with distinct priorities are designed as the initial solutions for local searches. A premium group includes  $NS_t$ . An alternative group contains  $|NS_t|$  solutions that are randomly selected from the rest population of generation  $t$  except  $NS_t$ . The rationale for this prioritized and equally-sized grouping is that it is more promising to exploit  $NS_t$  and exploitation on dominated solutions may waste computation resources. The alternative group is only used when a local searcher cannot find any qualified local optimum based on the premium group. The alternative group thus introduces randomness and adaptively raises the frequency of refinement.

If the CTS cannot find a qualified local optimum based on both premium and alternative groups, the entire search is considered stagnated and the stagnation time ( $T_{stag}$ ) accordingly increases by one. The entire search terminates if  $T_{stag}$  reaches

TABLE II  
POWER PROFILE AND REQUIRED LABOR OF AN EXTRUSION BLOW MOLDING MACHINE

State	Power (kW)	Duration (s)	Required personnel type
Off	0	$\geq 0$	None
Startup	3.51	442	Operator
Idle	1.19	$\geq 0$	
Preheat	17.52	1395	
PreheatIdle	8.15	$\geq 0$	
Proheat	16.95	810	Operator for powering up Technician for a changeover
ProheatIdle	9.00	$\geq 0$	
Production	46.35	17.92	Operator, technician, packer, and quality checker

the preset maximal stagnation time ( $T_{max}$ ) or the time budget is used up. Thereby,  $T_{max}$  may terminate an AMOMA instance before a time budget is used up, which provides room for an AMOMA instance to go even faster than the expectation.

IV. RESULTS AND DISCUSSIONS

The performance of AMOMA was validated in a case study of an EBM process in a plastic bottle manufacturer.

A. Case Study: Extrusion Blow Molding Process

The EBM machine comprises three major electricity consumers: main system, hydraulic system, and extruder. A Siemens PAC3200 power meter was installed on each of these consumers. Its sampling interval was 30 sec. Power data was sampled collected by a PLC (programmable logic controller) using the Modbus protocol. Fig. 4 shows the power data, where eight power states are identified: Off, Startup, Idle, Preheat, PreheatIdle, Proheat, ProheatIdle, and Production.

Consequently, the energy model of this EBM machine was established (Fig. 5). The operations above the transitional arrow in Fig. 5 should be performed by operators. When the machine is powered on, it goes through Startup, Idle, and Preheat, during which the plastic is heated in the barrel until 140 °C. It then stays at PreheatIdle and remains this temperature until an operator launches Proheat. Afterwards, the temperature of plastic rises between 140 °C and 200 °C, depending on the type of bottles to be produced. When the target temperature is achieved, the machine stays at ProheatIdle. Once a production command is given, it transitions to Production state.

The power profiles of this EBM machine and required labor per state are shown in Table II. The power profile of Production state does not depend on distinct types of bottles, as standard deviations of power and cycle time for producing different types of bottles only occupy 2% and 1% of the corresponding average. A changeover is performed at ProheatIdle state and has an average duration of 13309 sec. A workday comprises early shift (6 AM - 2 PM), late shift (2 PM - 10 PM), and night shift (10 PM - 6 AM). The labor compensation of a night shift rises by 10%. The factory is closed on weekends.

Note that the purpose of the following experiments is to demonstrate the effectiveness and efficiency of the AMOMA under time-varying electricity prices and labor wage, even with a large number of scheduling time slots. The two-week historical RTP data was thereby taken from the local electricity spot market, though this hourly-dynamic electricity price is known only one day in advance in reality. For real applications,

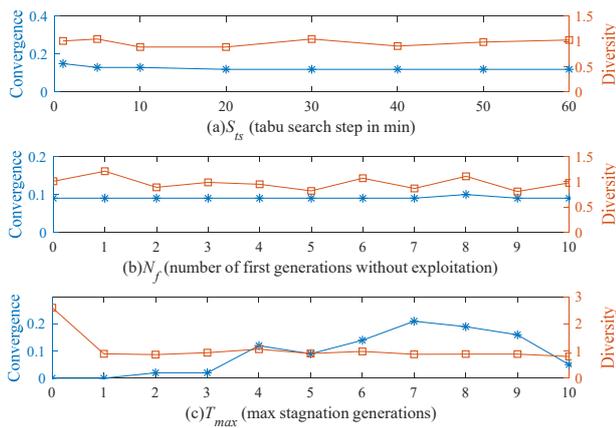


Fig. 6. Parametric sensitivity of the two proposed tabu searches (CTS and DTS) in convergence and diversity. A higher convergence rate and a lower diversity rate indicate superior Pareto front approximation.

the AMOMA can be used to produce a 24-h schedule under RTP. Artificial neural network (ANN)-based price forecasting may be used to enlarge the scheduling time span [24]. The scheduling time slot  $\delta t$  was set as 1 sec, leading to overall 1,209,600  $\delta t$ . Overall 10 independent jobs remained to be processed, whose duration varies from 8960 sec (500 bottles) to 71,680 sec (4000 bottles).

### B. Parameter Tuning of AMOMA

The AMOMA was run on a computer with Intel Core i5-3470 @ 3.2 GHz and 8 G RAM. The time budget was fixed at 2 min for each run. While  $\Delta$  was employed to indicate the diversity driven by parameter vector  $\vec{p}$ , the metric  $\Upsilon(\vec{p})$  was used to measure the convergence steered by  $\vec{p}$ :

$$\Upsilon(\vec{p}) = |NS^*(\vec{p})| / |NS(\vec{p})| \quad (19)$$

where  $NS^*(\vec{p})$  is the global approximation set (domination-based aggregation of all nondominated solutions) contributed by  $\vec{p}$ , and  $NS(\vec{p})$  are nondominated solutions produced by  $\vec{p}$ .  $\vec{p}$  is (population size, crossover rate, mutation rate) in tuning exploration and one of  $S_{ts}$ ,  $N_f$ , and  $T_{max}$  in turning exploitation. The AMOMA with each  $\vec{p}$  was independently run 50 times.

The NSGA-II was tuned for stronger convergence, without local searches. Following the tuning guidance in Sect. III-B, three promising levels for population size, crossover rate, and mutation rate were set as (100, 500, 1000), (0.7, 0.8, 0.9), and (0.1, 0.2, 0.3), respectively. Consequently, (1000, 0.9, 0.2) was selected due to its strongest convergence ( $\Upsilon$  was 16% compared to others between 0 and 11%).

Both CTS and DTS were further tuned for higher convergence and preserved diversity, with the former tuned NSGA-II. As indicated in Fig. 6a, the variation of  $S_{ts}$  within 1 h has little impact on both convergence and diversity. This is because the smallest time granule of electricity prices and labor wages is 1 h, such that a search step smaller than 1 h cannot enable finer exploitation in the solution space. The  $S_{ts}$  was thus set to 1 h to speed up exploitation without affecting the convergence and diversity. As shown in Fig. 6b, the pure

TABLE III  
PERFORMANCE COMPARISON (MEAN  $\pm$  STANDARD DEVIATION) OF SEVEN ALGORITHMS IN FIFTY RUNS USING THE EMPIRICAL DATA

Algorithm	$ NS ^a$	$\Upsilon^b$	$\Delta^c$
AMOMA	<b>10.28</b> $\pm$ 5.09	<b>0.33</b> $\pm$ 0.16	0.88 $\pm$ 0.56
NSGA-II	14.04 $\pm$ 2.51	0.05 $\pm$ 0.02	3.56 $\pm$ 0.35
GRASP	6.32 $\pm$ 1.82	0.14 $\pm$ 0.09	0.60 $\pm$ 0.35
MA-C	3.38 $\pm$ 0.67	0.06 $\pm$ 0.02	0.90 $\pm$ 0.49
MA-D	<b>9.28</b> $\pm$ 4.29	0.18 $\pm$ 0.10	0.87 $\pm$ 0.65
AMOMA-N	<b>9.76</b> $\pm$ 5.37	0.18 $\pm$ 0.12	0.87 $\pm$ 0.82
AMOMA-E	3.08 $\pm$ 1.73	0.06 $\pm$ 0.03	<b>0.30</b> $\pm$ 0.22

$|NS|^a$ : number of nondominated solutions,  $\Upsilon^b$ : convergence,  $\Delta^c$ : diversity

exploration in the first  $N_f$  generations moderately influences the diversity, while it has little impact on the convergence. This is explained by the global search characteristic of NSGA-II, such that it cannot guarantee exploitation in a potential region of a solution space to improve the convergence.  $N_f$  was thereby set to 2 to introduce local searches as early as possible while preserving the diversity. As implied in Fig. 6c, the convergence and diversity are sensitive to  $T_{max}$  and whether to go on upon stagnation. This reveals that the joint exploration and exploitation of the proposed AMOMA can effectively prevent premature convergence. Therefore,  $T_{max}$  was set to 7 in order to achieve superior levels in both convergence and diversity without having to terminate the search too early or too late.

### C. Scheduling of an Extrusion Blow Molding Process

1) *Benchmark*: Using the empirical data (Sect. IV-A) and tuned configurations (Sect. IV-B), the AMOMA was compared with NSGA-II [15], GRASP (greedy randomized adaptive search procedure) [25], MA-C (hybrid of NSGA-II and CTS), MA-D (hybrid of NSGA-II and DTS), AMOMA-N (AMOMA which only exploits nondominated solutions), and AMOMA-E (AMOMA which optimizes toward  $TEC$  and  $C_{max}$ ).

NSGA-II, MA-C, MA-D, AMOMA-N, and AMOMA-E remained the corresponding configurations and time budget for AMOMA. In the construction phase of GRASP, a solution was constructed by iteratively building a restricted candidate list (RCL) and randomly selecting a job for this solution. In the local search phase of GRASP, CTS and DTS were sequentially used as in AMOMA. The parameter  $\alpha$  (“ $\alpha$  percent” distant from the nondominated solutions) for building an RCL was tuned at 0:0.2:1. It was set as 1 with the highest  $\Upsilon$  (0.48), implying that full randomness effectively improved the convergence. Each algorithm was independently run 50 times.

The performance of an algorithm was evaluated in three dimensions: the number of nondominated solutions ( $|NS|$ ), the convergence ( $\Upsilon$ ), and the diversity ( $\Delta$ ). Intermediate  $|NS|$  is preferred, because small  $|NS|$  provides insufficient trade-off insights and large  $|NS|$  causes problems on optimal selection. High  $\Upsilon$  and low  $\Delta$  are preferred, respectively.

As presented in Table III, the AMOMA is among the best in  $|NS|$ , the best in convergence, and moderate in diversity. NSGA-II has the worst convergence by achieving the largest  $|NS|$  and smallest  $\Upsilon$ . Its diversity is also the worst, indicating that it is incapable to evenly diversify the convergence introduced by dispatching rules in the initial population. This highlights the need for exploitation. The relatively small  $|NS|$  of GRASP indicates its limitation in producing a set of

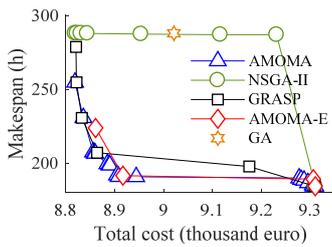


Fig. 7. Pareto front (PF) approximations using the empirical data.

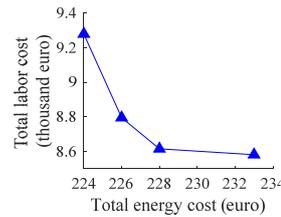


Fig. 8. Quantified trade-off between energy and labor costs.

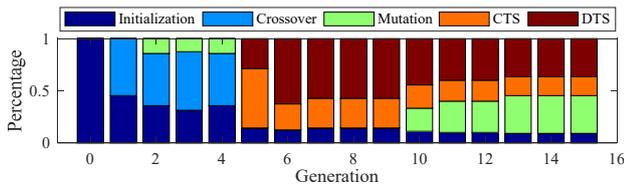


Fig. 9. Source of nondominated solutions of an AMOMA instance.

nondominated solutions, compared to the population-based AMOMA which better fits MOPs. The small  $|\text{NS}|$  of MA-C underlines the need of DTS to diversify the biased convergence introduced by the CTS. Pure CTS in MA-C cannot effectively enhance convergence. It needs the assist of DTS in diversity preservation to achieve comparable convergence of AMOMA. Although diversity preservation of DTS in MAC-D effectively strengthens the convergence compared to NSGA-II and MA-C, a lack of convergence enhancement measures prevents MA-D from achieving comparable convergence of AMOMA. The convergence of AMOMA-N (Table III) is nearly halved by exploiting only nondominated solutions (premium group). This emphasizes the contribution of the prioritized grouping strategy in AMOMA for convergence enhancement, where the alternative group introduces randomness for the search to escape a local optimum. The smallest  $|\text{NS}|$  and low  $\Upsilon$  of AMOMA-E reveal its incapability in minimizing  $TC$  due to its ignorance of  $TLC$ .

2) *Trade-off Analysis*: Fig. 7 shows the PF approximations of the best runs of AMOMA, NSGA-II, GRASP, and AMOMA-E regarding convergence, and a solution given by a genetic algorithm (GA) which optimizes toward  $TEC$ . Compared to NSGA-II, AMOMA is significantly more effective in reducing  $C_{max}$  while maintaining the large range of  $TC$ . The nearly constant large  $C_{max}$  of the solutions provided by NSGA-II implies its weak convergence for this proposed problem. It attempts to evidently shift production jobs over time to search for the minimization potential of  $TC$ . Consequently, despite its similar  $TC$  range compared to the AMOMA, this is evidently compensated by the prolonged  $C_{max}$ .

Although the PF approximations of GRASP and AMOMA-E are partially analogous to that of AMOMA, AMOMA remains superior in the entire range of  $TC$  (Fig. 7). An observation in the approximation set of AMOMA is that a slight prolongation of a short  $C_{max}$  (by 0.5%) can dramatically reduce  $TC$  (by 4%). However, a further decrease in  $TC$  (by 1%) has to be compromised by a significant rise in  $C_{max}$  (by 33%). This is because jobs tend to go across weekends to search for more economical periods when  $C_{max}$  rises. Nonetheless, the significantly raised time flexibility does not induce

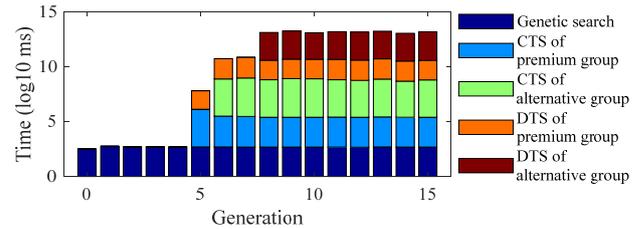


Fig. 10. CPU time consumed by major components of an AMOMA.

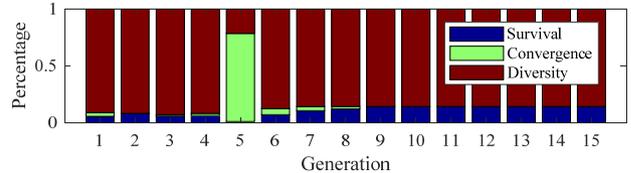


Fig. 11. Convergence, diversity, and survival rate of an AMOMA instance.

corresponding evident reduction in  $TC$ , since the RTP data is hourly dynamic but daily similar on weekdays and  $TLC$  dominates  $TEC$  in this case study.

The single solution provided by the GA is far away from the approximation set of AMOMA (Fig. 7). This proves that the existing energy-aware scheduling methods cannot optimize  $TC$  and  $C_{max}$ , which are important production metrics. Such a poor result is explained by two reasons. Firstly, analogous to the NSGA-II, the GA focuses on the global search without any guarantee on exploitation in the solution space. Secondly, the total labor cost cannot be explicitly minimized in this GA, such that the  $TC$  cannot be efficiently minimized, though this GA enlarges the  $C_{max}$  to search for more  $TC$  minimization potential.

The trade-off relationship of  $TEC$  and  $TLC$  is sketched in Fig. 8, which was obtained by setting objectives as  $TEC$  and  $TLC$ . A slight decrease in  $TEC$  from 228 to 224 euro (2%) can lead to an obvious increase in  $TLC$  from 8.6 to 9.3 thousand euro (8%). This could be because  $TEC$  is evidently smaller than  $TLC$  in this case study. Consequently,  $TLC$  is more sensitive to production load shifting over time, compared to  $TEC$ . This again highlights the economic necessity to jointly consider energy and labor in production load shifting.

3) *Adaptation Behavior*: The AMOMA will be proven to have a synergistic and adaptive balance between exploration and exploitation. Fig. 9 demonstrates the dynamic percentage of AMOMA operators in producing nondominated solutions in an AMOMA search. Genetic search operators (mem-integrated initialization, crossover, and mutation) retain full occupation in the first four generations. Afterward, local search operators (CTS and DTS) dominate while crossover remains a portion of zero. This reveals that exploration and exploitation stay strong and weak in the early stage of a search, respectively, and vice versa in the late stage. In the early stage, potential regions have to be extensively explored. In the late stage, these potential regions should be extensively exploited before terminating the entire search. This reasoning is further demonstrated in Fig. 10. The CPU time is evenly shared among genetic search and two local searches that are applied to two groups, respectively. This exhibits the balanced intensity of different operators. The search makes increasing and full use of

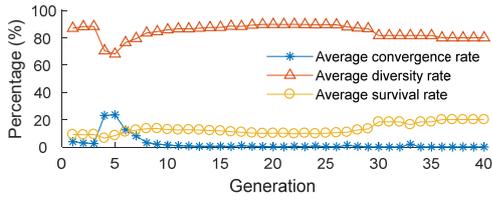


Fig. 12. Average convergence, diversity, and survival rates over 50 independent runs of the AMOMA.

exploitation in generations 5-7 and 8-15, respectively. This shows the adaptation behavior of AMOMA.

Inspired by the cross-dominance metric  $\lambda$  [23], the cross-dominance and equality metrics are introduced to measure the diversity and survival of  $NS_t$  compared with  $NS_{t-1}$ . Cross-dominance indicates the nondominance between solutions that have at least one different objective value. Equality means that two solutions are equal in all objective values. Fig. 11 presents a dynamic change of these metrics in an AMOMA instance. Throughout the search, the diversity remains high and the survival rate steadily rises. This implies that the premature convergence is avoided in AMOMA. Conversely, the convergence rate varies a lot. It is slightly contributed by the genetic search in generations 1-4 (0-3%), and is significantly contributed by the CTS in generation 5 (77%). It gradually reduces after generation 6 until falling to zero in generation 9 and remaining zero in generations 10-15. The single peak of convergence demonstrates the deepest descent local search property of CTS, which contributes to the fast convergence.

Fig. 12 further presents the average convergence, diversity, and survival rates over these 50 independent AMOMA instances. Note that the number of generations varies between 11 and 40 in these 50 runs, due to the self-adaptive termination criterion of AMOMA (Sect. III-D). Therefore, the calculation of average rates between generations 11 and 40 only considers the runs that has the corresponding generation number. As clearly exhibited by Fig. 12, the convergence curve nearly gets saturated even for the smallest generation number (11). This indicates the fast convergence of AMOMA within only 2 min, satisfying industrial DR's or production scheduling's requirement of fast and high-quality decision making. Besides, the peak average convergence rate at around generation 5 exhibits that the premature convergence can be generally avoided. The average trends of these three rates comply with these shown in Fig. 11. This implies that the former analyzed adaptation behavior of AMOMA is representative.

4) *Economic Sensitivity*: In the former case study,  $TEC$  is dominated by  $TLC$  and production is not allowed on weekends. Two other cases were assumed by scaling the power of EBM machine:  $TEC$  is comparable to  $TLC$  (power scale was 50) and  $TEC$  dominates  $TLC$  (power scale was 1000). Each case had two scenarios: enabled and disabled weekend production. The AMOMA with the former configuration was repeated 50 times for each scenario. The normalized range of  $TC$  and  $C_{max}$  of an approximation set was obtained by Eqs. (20-21), respectively:

$$r(TC) = [\max(TC) - \min(TC)] / \max(TC) \quad (20)$$

$$r(C_{max}) = [\max(C_{max}) - \min(C_{max})] / \min(C_{max}) \quad (21)$$

where  $r(TC)$  characterizes the maximal  $TC$  reduction potential of AMOMA given a 2-min time budget, and  $r(C_{max})$  measures

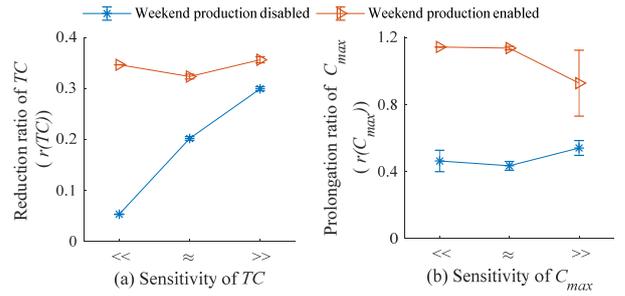


Fig. 13. Statistical reduction potential of  $TC$  and prolongation trend of  $C_{max}$  (mean and standard deviation) in three scenarios (<<:  $TEC$  is dominated by  $TLC$ ,  $\approx$ :  $TEC$  is comparable to  $TLC$ , >>:  $TEC$  dominates  $TLC$ ).

the maximal prolongation rate of  $C_{max}$  that has to be compensated to achieve this  $r(TC)$ .

As indicated by Fig. 13a, the option of weekend production impacts  $r(TC)$ . If it is enabled,  $r(TC)$  remains a relatively high level and is nearly insensitive to the proportion of  $TEC$  and  $TLC$ . This is because weekends provide the AMOMA more periods with lower electricity prices to optimize  $TC$  considering the trade-off between  $TEC$  and  $TLC$ . When weekend production is disabled,  $r(TC)$  increases with the rising share of  $TEC$  in  $TC$  (Fig. 13a). This is explained by the higher-priced periods on weekdays, such that the increasing portion of  $TEC$  in  $TC$  makes  $r(TC)$  much more sensitive to load shifting over time.

Fig. 13b demonstrates that if  $TEC$  does not dominate  $TLC$ ,  $r(TC)$  stays constant regardless of weekend production. This implies that if the share of  $TEC$  is not significant, the  $C_{max}$  will increase by about 120% and 40% to achieve the  $TC$  reduction ratio in Fig. 13a with and without weekend production, respectively. If  $TEC$  dominates  $TLC$ ,  $r(TC)$  without weekend production slightly grows, as higher time flexibility is needed for more lower-priced periods to reduce the dominant  $TEC$  and thus  $TC$ ; comparatively,  $r(TC)$  with weekend production moderately drops with a larger variation, implying that the major  $TEC$  portion in  $TC$  and the increased lower-pricing periods reduce the  $C_{max}$  prolongation for  $TC$  minimization despite the rising variation.

#### D. Generalization of the Proposed Scheduling Method

Although this paper focuses on single-machine scheduling, the proposed method to model and integrate energy and labor-awareness in conventional production scheduling algorithms can be easily extended to other shop floor configurations. The following experiment takes the partial flexible job shop (PFJS), one of the most complex shop floor configurations, as an illustration of such an extension.

In this experiment, the  $8 \times 8 \times 27$  (number of machines  $\times$  number of jobs  $\times$  number of operations) PFJS scheduling problem instance was taken from the common benchmark instance set [26]. The energy and labor data remained these in the previous experiment. Five objectives were involved:  $C_{max}$ ,  $TEC$ ,  $TLC$ , maximum workload ( $MWL$ ), and total workload ( $TWL$ ). Compared to the previous experiment, the calculation of  $C_{max}$ ,  $TEC$  and  $TLC$  considered the aggregation effect of multiple machines.  $MWL$  indicates the maximum work time spent on any of the involved machines.  $TWL$  measures the total

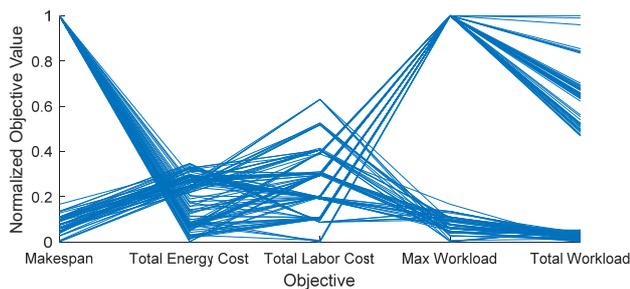


Fig. 14. Quantified relation among the makespan, the total energy cost, the total labor cost, the maximum workload, and the total work load for the energy- and labor-aware production scheduling of multiple machines under time-varying electricity prices and labor wage.

work time of all machines. NSGA-III [27] was taken as the genetic search-based exploration due to the number of objectives larger than 3. The time budget was 3 min.

Fig. 14 demonstrates the obtained nondominated solutions on the parallel coordinates with their normalized objective values. In this plot, a curve across the 5 objectives represents a nondominated solution. The superposition of all curves reveals the relation among these 5 objectives. As exhibited in Fig. 14,  $C_{max}$  strongly conflicts with  $TEC$ , which complies with the relation quantified in the previous experiment. This is analogously explained by the shift effect of this scheduling method on the time horizon, such that longer free durations are created to search for the lower-priced periods for  $TEC$  minimization. Similar to the observation in Fig. 8,  $TEC$  shows a trade-off relation with  $TLC$  in Fig. 14, though this relation is relatively weak. This is analogously due to the contradicted electricity prices and labor wage over time. Due to the high values of  $C_{max}$  in many nondominated solutions, it could be inferred that the trade-off between  $C_{max}$  and  $TC$  also remains.  $TLC$  has a notable trade-off relation with  $MWL$  (Fig. 14). If  $MWL$  is reduced, jobs are more evenly distributed on machines, such that more labor shifts are likely to be triggered resulting increased  $TLC$ . Conversely, if  $MWL$  rises, the work load and labor shifts are more concentrated on a machine, such that less labor shifts and lower  $TLC$  are incurred. Nevertheless,  $MWL$  and  $TWL$  have a nearly harmonious relation (Fig. 14). The lower ranges of  $TEC$  and  $TLC$  additionally demonstrates the effectiveness of the proposed energy- and labor-aware scheduling method in simultaneous minimization of both cost parts while considering their trade-off relation.

## V. SCHEDULING CONCLUSION

This paper presented a multiobjective energy- and labor-aware production scheduling model under real-time electricity pricing, and AMOMA to solve this model. This model minimizes makespan and joint energy and labor cost under constraints of price-aware energy consumption, labor, job, changeover, and machine. As two local search operators for convergence enhancement and diversity preservation, respectively, CTS and DTS were reactively launched upon a cross-dominance-based convergence rate of zero. Besides a premium group for local searches, an ordinary group was used

to raise the refinement frequency when no qualified local optimum was found from the former group. This prioritized grouping strategy was demonstrated to evidently strengthen convergence. With only a 2-min time budget and an ordinary computer, an AMOMA instance could stop even earlier, comparing the monitored and maximal stagnation times. Through a case study of an extrusion blow molding machine and extensive performance comparison with existing algorithms, these hybridization and adaptation measures were proven effective to achieve fast Pareto front convergence without deteriorating the diversity. The generalization of this scheduling method was further demonstrated in a multi-machine shop floor configuration. The further work will investigate the large-scale optimization regarding the number of time slots, jobs, operations, machines, and human workers.

## REFERENCES

- [1] T. Strasser, F. Andr n, J. Kathan, C. Cecati, C. Buccella, P. Siano, P. Leit o, G. Zhabelova, V. Vyatkin, P. Vrba, and V. Mařík, "A Review of Architectures and Concepts for Intelligence in Future Electric Energy Systems," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2424-2438, 2015.
- [2] L. Park, Y. Jang, S. Cho, and J. Kim, "Residential Demand Response for Renewable Energy Resources in Smart Grid Systems," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 3165-3173, 2017.
- [3] N. G. Paterakis, O. Erdinç, A. G. Bakirtzis, and J. P. S. Catal o, "Optimal Household Appliances Scheduling Under Day-Ahead Pricing and Load-Shaping Demand Response Strategies," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 6, pp. 1509-1519, 2015.
- [4] C. Gahn, F. Denz, M. Dirr, and A. Tuma, "Energy-efficient scheduling in manufacturing companies: A review and research framework," *European Journal of Operational Research*, vol. 248, no. 3, pp. 744-757, 2016.
- [5] Y. M. Ding, S. H. Hong, and X. H. Li, "A Demand Response Energy Management Scheme for Industrial Facilities in Smart Grid," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2257-2269, 2014.
- [6] X. Gong, T. De Pessemer, W. Joseph, and L. Martens, "A generic method for energy-efficient and energy-cost-effective production at the unit process level," *Journal of Cleaner Production*, vol. 113, pp. 508-522, 2016.
- [7] V. V. Unhelkar, P. A. Lasota, Q. Tyroller, R. D. Buhai, L. Marceau, B. Deml, and J. A. Shah, "Human-Aware Robotic Assistant for Collaborative Assembly: Integrating Human Motion Prediction with Planning in Time," *IEEE Robotics and Automation Letters*, vol. PP, no. 99, pp. 1-1, 2018.
- [8] Y. Wang, and L. Li, "Time-of-use based electricity cost of manufacturing systems: Modeling and monotonicity analysis," *International Journal of Production Economics*, vol. 156, pp. 246-259, 2014.
- [9] Y. Wang, and L. Li, "Manufacturing profit maximization under time-varying electricity and labor pricing," *Computers & Industrial Engineering*, vol. 104, pp. 23-34, 2017.
- [10] X. Gong, M. Van der Wee, T. De Pessemer, S. Verbrugge, D. Colle, L. Martens, and W. Joseph, "Integrating labor awareness to energy-efficient production scheduling under real-time electricity pricing: An empirical study," *Journal of Cleaner Production*, vol. 168, pp. 239-253, 2017.
- [11] W.-a. Yang, Y. Guo, and W. Liao, "Multi-objective optimization of multi-pass face milling using particle swarm intelligence," *The International Journal of Advanced Manufacturing Technology*, vol. 56, no. 5, pp. 429-443, 2011.
- [12] M. Gen, W. Zhang, L. Lin, and Y. Yun, "Recent advances in hybrid evolutionary algorithms for multiobjective manufacturing scheduling," *Computers & Industrial Engineering*, 2017.
- [13] M. L. Pinedo, "General purpose procedures for deterministic scheduling," *Scheduling: theory, algorithms, and systems*, pp. 375-393: Springer International Publishing, 2016.
- [14] X. D. Xue, K. W. E. Cheng, T. W. Ng, and N. C. Cheung, "Multi-Objective Optimization Design of In-Wheel Switched Reluctance Motors

- in Electric Vehicles," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 9, pp. 2980-2987, 2010.
- [15] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, 2002.
- [16] L. A. Pereira, S. Haffner, G. Nicol, and T. F. Dias, "Multiobjective Optimization of Five-Phase Induction Machines based on NSGA-II," *IEEE Transactions on Industrial Electronics*, vol. PP, no. 99, pp. 1-1, 2017.
- [17] K. Sindhya, K. Miettinen, and K. Deb, "A Hybrid Framework for Evolutionary Multi-Objective Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 4, pp. 495-511, 2013.
- [18] F. Neri, and C. Cotta, "Memetic algorithms and memetic computing optimization: A literature review," *Swarm and Evolutionary Computation*, vol. 2, pp. 1-14, 2012.
- [19] X. Chen, Y. S. Ong, M. H. Lim, and K. C. Tan, "A Multi-Facet Survey on Memetic Computation," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 5, pp. 591-607, 2011.
- [20] A. Maesani, G. Iacca, and D. Floreano, "Memetic Viability Evolution for Constrained Optimization," *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 1, pp. 125-144, 2016.
- [21] T. Aouam, K. Geryl, K. Kumar, and N. Brahimi, "Production planning with order acceptance and demand uncertainty," *Computers & Operations Research*, vol. 91, pp. 145-159, 2018.
- [22] F. Qiao, Y. Ma, M. Zhou, and Q. Wu, "A Novel Rescheduling Method for Dynamic Semiconductor Manufacturing Systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. PP, no. 99, pp. 1-11, 2018.
- [23] A. Caponio, and F. Neri, "Integrating Cross-Dominance Adaptation in Multi-Objective Memetic Algorithms," *Multi-Objective Memetic Algorithms*, C.-K. Goh, Y.-S. Ong and K. C. Tan, eds., pp. 325-351, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009.
- [24] X. Huang, S. H. Hong, and Y. Li, "Hour-Ahead Price Based Energy Management Scheme for Industrial Facilities," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 6, pp. 2886-2898, 2017.
- [25] G. Mouzon, and M. B. Yildirim, "A framework to minimise total energy consumption and total tardiness on a single machine," *International Journal of Sustainable Engineering*, vol. 1, no. 2, pp. 105-116, 2008.
- [26] I. Kacem, S. Hammadi, and P. Borne, "Approach by localization and multiobjective evolutionary optimization for flexible job-shop scheduling problems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 32, no. 1, pp. 1-13, 2002.
- [27] K. Deb, and H. Jain, "An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems With Box Constraints," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 4, pp. 577-601, 2014.