

Optimal control of inventory-production systems in presence of perturbations on the inventory level

Davide Giglio and Silvia Siri

Abstract—A set of optimal control strategies is determined in this paper for a class of inventory-production systems, in order to optimize the production process in presence of perturbations on the inventory level. The inventory-production system has to satisfy a time-varying external demand, which is assumed to be a piecewise constant function of time; the optimization involves both production speed and the time instants at which the speed is changed. The inventory level, which is the state of the system, is supposed to be subject to some perturbations such that, at some time instants, it can be lower or higher than expected. Starting from the optimal (open-loop) solution found by the authors in a previous work, in this paper the optimal production speed is determined as an optimal (closed-loop) control strategy, which is a function of the system state. Such strategy minimizes the extra-cost that is paid by the system in consequence of a perturbation.

I. INTRODUCTION

In this paper a manufacturing system is considered as a single-stage process in which raw parts are transformed into processed parts; the decisions regard the pattern of the production rate, in order to meet a deterministic external demand expressed as a positive piecewise constant function which changes value at asynchronous time instants. This system can be related to the class of lot sizing models, in which the main decisions concern when and how much to produce in order to minimize setup, production, and holding costs [1], [2]. The first research work on lot sizing models was developed in the fifties [3], for the uncapacitated case and, since then, this class of models have been studied by researchers both in the operations research field and in the optimal control community [4]. Analyzing the different approaches present in the literature on lot sizing problems, both continuous-time [5], [6] and discrete-time [7], [8] models can be found, together with few works in which the time intervals in the planning horizon have different lengths [9]. The main peculiarity of the inventory-production model adopted in this paper stands in considering the external demand and the production effort as piecewise constant functions whose values change in correspondence with asynchronous time instants. The decisions concern not only the values of the production rate over time, but also the time instants at which these values change.

The objective of this paper is to find optimal control strategies for the considered class of inventory-production systems. In the literature, different approaches exist for the optimal control of manufacturing systems. An important

research field in this area regards failure-prone production systems represented with fluid models in a continuous-time framework [10]–[12], where failure/repair stochastic processes are considered. Other approaches of optimal control of production systems can be found in the area of scheduling models, corresponding to discrete-event systems whose dynamics is driven by specific events (such as the arrivals of jobs) that are asynchronous in time [13]–[16]. As in this latter research field, also in the model adopted in this paper asynchronous time instants characterize the dynamics of the system; however the production system is here considered at a more aggregate level.

The class of inventory-production systems considered in this paper has been dealt with in [17] by the same authors. In that work, a planning problem for determining the optimal pattern of the production rate has been formulated, some properties of the optimal solution of the problem have been proven, and an algorithm providing the optimal solution in polynomial time has been proposed. Objective of the present work is to exploit the results provided in [17], in order to find a set of optimal (closed-loop) control strategies as functions of the system state. Closed-loop strategies are useful or even necessary when the system is affected by perturbations. It is here supposed that the system state (that is, the inventory level) is measured at some generic time instants; after these measurements, it can be verified that the parts contained in the inventory are less or more than expected, because of a perturbation occurred before (such as a demand different from the expected one for some time periods or malfunctions to the power system or the machines, leading to a production effort different from the expected one). The optimal control strategies are defined with reference to the optimal path of the production rate (that is, without any perturbations), in order to minimize the extra-cost paid by the system in consequence of a perturbation.

This paper is organized as follows. In the next section, the optimization problem for the considered class of systems is stated and the optimal solution is provided. Some considerations about the perturbations which may affect the system are made in Section III, whereas in Section IV the optimal control strategies are provided in connection with the two cases in which the inventory contains less or more parts than expected. Some conclusions are drawn in Section V.

II. OPTIMIZATION OF A CLASS OF INVENTORY-PRODUCTION SYSTEMS

The considered model of inventory-production system, the relevant optimization problem and the main results on the

D. Giglio and S. Siri are with the Department of Computer Science, Bioengineering, Robotics and Systems Engineering (DIBRIS), University of Genova, Via Opera Pia 13, 16145 – Genova, Italy
davide.giglio@unige.it silvia.siri@unige.it

structural properties of any optimal solution are concisely reported in this section. The reader can find more details and the formal proofs of the theorems in [17].

A. The system model

In the considered inventory-production system a production resource transforms raw parts into processed parts, that can be stored in an inventory, in order to be delivered outside according to a positive external demand to be satisfied. $x(t)$, $u(t)$ and $e(t)$ denote the inventory level, the production speed (or rate), and the external demand, respectively, at time t .

The external demand is assumed to be a piecewise constant function of time defined in the interval $[0, T]$: in particular, $e_k > 0$, $k = 1, \dots, Q$, is the rate of products required in time interval $[\theta_{k-1}, \theta_k)$ and θ_k , $k = 1, \dots, Q - 1$, is the asynchronous time instant at which this rate changes, with $\theta_0 = 0$ and $\theta_Q = T$. Analogously, the production speed is modeled as a piece-wise constant function of t ; in this connection, $0 \leq u_h \leq U$, $h = 1, \dots, P$, is the speed within time interval $[\tau_{h-1}, \tau_h)$, U is the maximum allowed speed, and τ_h , $h = 1, \dots, P - 1$, is the asynchronous time instant in which this rate changes, with $\tau_0 = 0$ and $\tau_P = T$. The state of the system is represented by the inventory level $x(t)$, which has a piecewise linear pattern.

For this class of inventory-production systems, a finite-horizon optimization problem is stated, whose objective is to find the optimal production speed over time in order to minimize setup, production, and holding costs. The decision variables are the time instants τ_h , $h = 1, \dots, P - 1$, the production efforts u_h , $h = 1, \dots, P$, and the number of "discontinuity points" P .

B. The optimization problem

The setup, production and holding costs are, respectively, given by

$$C_S = \sum_{h=0}^{P-1} \sigma y_h \quad (1)$$

$$C_P = \psi \sum_{h=1}^P u_h (\tau_h - \tau_{h-1}) \quad (2)$$

$$\begin{aligned} C_H = \varphi \int_0^T x(t) dt = \varphi \left[x_0 T + \sum_{h=1}^P \frac{1}{2} u_h (\tau_h - \tau_{h-1})^2 \right. \\ \left. + \sum_{h=2}^P \sum_{p=1}^{h-1} u_p (\tau_p - \tau_{p-1}) (\tau_h - \tau_{h-1}) \right. \\ \left. - \sum_{k=1}^Q \frac{1}{2} e_k (\theta_k - \theta_{k-1})^2 \right. \\ \left. - \sum_{k=2}^Q \sum_{q=1}^{k-1} e_q (\theta_q - \theta_{q-1}) (\theta_k - \theta_{k-1}) \right] \quad (3) \end{aligned}$$

being σ the setup cost to be paid when the production resource passes from an idle to a working state, ψ the unitary processing cost, φ the unitary holding cost, and $x(0) = x_0 \geq 0$ the initial inventory level. In (1), y_h is a binary variable whose value is equal to 1 if a setup is required at τ_h , $h = 0, \dots, P - 1$ (it is assumed that the resource is initially idle and no setup is required at the end of the production process).

In order to impose that the inventory is non negative at each time instant, it is sufficient to verify this condition in the asynchronous time instants τ_h , $h = 1, \dots, P$, and θ_k , $k = 1, \dots, Q$, where the slope of the inventory changes. In

order to write the expression of the inventory level in these time instants, it is necessary to introduce a set of binary variables $\rho_{h,k}$ defining for any pair (h, k) , $h = 0, \dots, P$, and $k = 0, \dots, Q$, the relative position between τ_h and θ_k (specifically, $\rho_{h,k}$ is equal to 0 if $\tau_h \leq \theta_k$ and 1 otherwise).

The optimization problem for the considered inventory-production system is defined as follows.

Problem 1: Given the initial inventory level $x_0 \geq 0$, the maximum production speed U , the time horizon T , and the external demand expressed by θ_k , $k = 0, \dots, Q$, with $\theta_0 = 0$ and $\theta_Q = T$, and e_k , $k = 1, \dots, Q$, find

$$\min_{\substack{\tau_h, h=1, \dots, P-1 \\ u_h, h=1, \dots, P \\ y_h, h=0, \dots, P-1 \\ \rho_{h,k}, h=1, \dots, P-1, k=1, \dots, Q-1}} C_S + C_P + C_H \quad (4)$$

being C_S , C_P , and C_H provided by (1), (2), and (3), respectively, with $\tau_0 = 0$, $\tau_P = T$, $\rho_{0,k} = 0$, $k = 0, \dots, Q$, $\rho_{h,0} = 1$, $h = 1, \dots, P$, $\rho_{P,k} = 1$, $k = 1, \dots, Q - 1$, $\rho_{h,Q} = 0$, $h = 1, \dots, P$, subject to

$$\begin{aligned} x_0 + \sum_{p=1}^h u_p (\tau_p - \tau_{p-1}) - \sum_{k=1}^Q \rho_{h,k} e_k (\theta_k - \theta_{k-1}) \\ - \sum_{k=1}^Q (\rho_{h,k-1} - \rho_{h,k}) e_k (\tau_h - \theta_{k-1}) \geq 0, \\ h = 1, \dots, P \quad (5) \end{aligned}$$

$$\begin{aligned} x_0 + \sum_{h=1}^P (1 - \rho_{h,k}) u_h (\tau_h - \tau_{h-1}) \\ + \sum_{h=1}^P (\rho_{h,k} - \rho_{h-1,k}) u_h (\theta_k - \tau_{h-1}) \\ - \sum_{q=1}^k e_q (\theta_q - \theta_{q-1}) \geq 0, \quad k = 1, \dots, Q \quad (6) \end{aligned}$$

$$\tau_h > \tau_{h-1}, \quad h = 1, \dots, P \quad (7)$$

$$0 \leq u_h \leq U, \quad h = 1, \dots, P \quad (8)$$

$$\tau_h - \theta_k + T(1 - \rho_{h,k}) > 0, \quad h = 1, \dots, P - 1 \\ k = 1, \dots, Q - 1 \quad (9)$$

$$\theta_k - \tau_h + T \rho_{h,k} \geq 0, \quad h = 1, \dots, P - 1 \\ k = 1, \dots, Q - 1 \quad (10)$$

$$|u_h - u_{h+1}| > 0, \quad h = 1, \dots, P - 1 \quad (11)$$

$$u_1 + U(1 - y_0) > 0 \quad (12)$$

$$-u_1 + U y_0 \geq 0 \quad (13)$$

$$u_h + U y_h > 0, \quad h = 1, \dots, P - 1 \quad (14)$$

$$-u_h + U(1 - y_h) \geq 0, \quad h = 1, \dots, P - 1 \quad (15)$$

$$P \in \mathbb{N}_{>0} \quad (16)$$

$$\rho_{h,k} \in \{0, 1\}, \quad h = 1, \dots, P - 1 \\ k = 1, \dots, Q - 1 \quad (17)$$

$$y_h \in \{0, 1\}, \quad h = 0, \dots, P - 1 \quad (18)$$

In Problem 1, constraints (5) and (6) impose the non-negativity of the inventory level, constraints (7) impose that the time instants in which the production effort changes are sequential, whereas constraints (8) fix the lower and upper bounds for the production effort. Moreover, constraints (9) and (10) ensure the correct values to binary variables $\rho_{h,k}$, constraints (11) impose that consecutive production rates are different, and, finally, constraints (12)-(15) impose the correct values for binary variables y_h .

C. The optimal solution

Problem 1 has a parametric structure, since P belongs to the set of decision variables, and it includes non linearities and combinatorial aspects. Anyway, it is possible to prove some results allowing to find analytically the optimal solution of the problem.

A first result on the optimal solution of Problem 1 states that the number of parts to be produced in $[0, T]$ is $n_P = \max\{0, \sum_{k=1}^Q e_k(\theta_k - \theta_{k-1}) - x_0\}$; as a matter of fact, in this problem, there is no advantage in producing more parts than those which are necessary to satisfy exactly the external demand and it is not possible to produce less parts than the necessary since the external demand must be completely satisfied (no negative inventory is allowed). Then, the optimal number of parts to be produced is always equal to the total demand minus the initial inventory; of course, no parts are produced in case the initial inventory is higher than the total demand.

Another important result concerns the presence of idle periods in an optimal solution. It can be proven that, in any optimal solution, there is one idle period at most; when the idle period is present, it is at the beginning of the considered time horizon. This statement can be explained by showing that an idle period in the middle of the production process would lead to higher holding and setup costs. It is also possible to write an analytical expression for the time instant τ_1 at which the idle period (if present) ends; this result is achieved by considering that the idle period, if present, lasts as much as possible, i.e. τ_1 is the farthest (from 0) time instant for which the solution remains feasible.

As regards the production, it can be shown that, in any optimal solution of Problem 1, the production speed in the working period is either equal to e_k or U . This is due to the fact that the optimal pattern of the production speed follows the pattern of the external demand, except when it is necessary to produce parts in advance, because of the maximum limit on the production speed. Then, if $e_k \geq U$, the production speed is equal to U ; otherwise, if $e_k < U$, the production speed can be either equal to e_k , if it is necessary only to satisfy the present demand, or equal to U , if it must satisfy some future demands.

On the basis of such considerations, the optimal values of production speed are provided by the following theorems, in which ξ_k ("residual stock"), $k = 0, \dots, Q$, represents the portion of the initial inventory level which is still available at time instant θ_k , having satisfied the external demands until that time instant, and X_k ("safety stock") represents the number of parts that the production resource must guarantee at θ_k in order to satisfy the future demands, from θ_k onwards.

Theorem 1: The optimal production speed in the generic interval of the working period, namely $[\theta_{k-1}, \theta_k)$, for any $k \in \{1, \dots, Q\}$ such that $(\xi_{k-1} \leq 0) \vee (\xi_{k-1} = X_{k-1})$, is equal to:

- in case $(X_{k-1} = 0) \wedge (X_k = 0)$:

$$u_{h+1} = e_k \quad (19)$$

and $\tau_h = \theta_{k-1}$ is the time instant at which the production speed changes to e_k (when $k \geq 2$) or the time instant at which the production starts (when $k = 1$);

- in case $(X_{k-1} = 0) \wedge (0 < X_k < (U - e_k)(\theta_k - \theta_{k-1}))$:
$$\begin{cases} u_{h+1} = e_k & \text{in } [\theta_{k-1}, \tau_{h+1}) \\ u_{h+2} = U & \text{in } [\tau_{h+1}, \theta_k) \end{cases} \quad (20)$$

and $\tau_h = \theta_{k-1}$ is the time instant at which the production speed changes to e_k (when $k \geq 2$) or the time instant at which the production starts (when $k = 1$), whereas $\tau_{h+1} = \theta_k - \frac{X_k}{U - e_k}$ is the time instant at which the production speed changes to U ;

- in case $(X_{k-1} > 0)$ or in case $(X_k = (U - e_k)(\theta_k - \theta_{k-1}))$:

$$u_{h+1} = U \quad (21)$$

and $\tau_h = \theta_{k-1}$ is the time instant at which the production speed changes to U (when $k \geq 2$ and if $(X_k = (U - e_k)(\theta_k - \theta_{k-1})) \wedge (e_{k-1} < U)$; otherwise no change of production speeds occurs within the considered interval) or the time instant at which the production starts (when $k = 1$).

being $h \in \{0, \dots, P - 1\}$ the number of production speed changes occurring before θ_{k-1} .

Theorem 2: The optimal production speed in the time interval in which the idle period (if present) ends, namely $[\theta_{k-1}, \theta_k)$, with $k \in \{1, \dots, Q\}$ such that $[(\xi_{k-1} > 0) \wedge (X_{k-1} < \xi_{k-1}) \wedge (\xi_k < 0)] \vee [(X_{k-1} < \xi_{k-1}) \wedge (X_k > \xi_k) \wedge (\xi_k > 0)]$, is equal to:

- in case $(\xi_k < 0) \wedge (X_{k-1} = 0) \wedge (X_k = 0)$:

$$\begin{cases} u_1 = 0 & \text{in } [\theta_{k-1}, \tau_1) \\ u_2 = e_k & \text{in } [\tau_1, \theta_k) \end{cases} \quad (22)$$

and $\tau_1 = \theta_{k-1} + \frac{\xi_{k-1}}{e_k}$ is the time instant at which the production starts;

- in case $(\xi_k < 0) \wedge (X_{k-1} = 0) \wedge (X_k > 0) \wedge (\frac{X_k}{U - e_k} < \frac{-\xi_k}{e_k})$:

$$\begin{cases} u_1 = 0 & \text{in } [\theta_{k-1}, \tau_1) \\ u_2 = e_k & \text{in } [\tau_1, \tau_2) \\ u_3 = U & \text{in } [\tau_2, \theta_k) \end{cases} \quad (23)$$

and $\tau_1 = \theta_{k-1} + \frac{\xi_{k-1}}{e_k}$ is the time instant at which the production starts, whereas $\tau_2 = \theta_k - \frac{X_k}{U - e_k}$ is the time instant at which the production speed changes to U ;

- in case $(\xi_k < 0) \wedge (X_{k-1} = 0) \wedge (X_k > 0) \wedge (\frac{X_k}{U - e_k} \geq \frac{-\xi_k}{e_k})$ or in case $(X_{k-1} < \xi_{k-1}) \wedge (\xi_k < 0) \wedge (X_{k-1} > 0)$ or in case $(\xi_k > 0)$:

$$\begin{cases} u_1 = 0 & \text{in } [\theta_{k-1}, \tau_1) \\ u_2 = U & \text{in } [\tau_1, \theta_k) \end{cases} \quad (24)$$

and $\tau_1 = \theta_k - \frac{X_k - \xi_k}{U}$ is the time instant at which the production starts.

These theorems allow determining in practice the optimal values of the decision variables of Problem 1. As a matter of fact, an optimal solution to Problem 1 can be obtained by, first of all, computing the residual stocks ξ_k , $k = 0, \dots, Q$ (through the forward relation $\xi_k = \xi_{k-1} - e_k(\theta_k - \theta_{k-1})$,

with $\xi_0 = x_0$) and the safety stocks X_k , $k = 0, \dots, Q$ (through the backward relation $X_{k-1} = \{0, X_k + (e_k - U)(\theta_k - \theta_{k-1})\}$, with $X_Q = 0$) and, secondly, implementing an algorithm that applies the results reported in Theorems 1 and 2, providing the optimal values of the decision variables P , τ_h , $h = 1, \dots, P-1$, u_h , $h = 1, \dots, P$, y_h , $h = 0, \dots, P-1$, and $\rho_{h,k}$, $h = 1, \dots, P-1$, $k = 1, \dots, Q-1$ (such an algorithm can be found in [17]).

D. Example – Part I

Consider an inventory-production system characterized by a maximum speed $U = 20$ parts/hour and an initial inventory level $x_0 = 17$ parts. The external demand $e(t)$ is the piecewise constant function illustrated in Fig. 1.

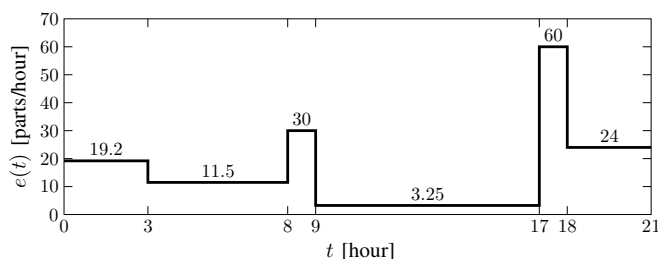


Fig. 1. Example – External demand $e(t)$.

The optimal production rate, obtained by solving Problem 1 through the results discussed in the previous subsection, is illustrated in Fig. 2. When such an optimal rate is employed, the pattern of the inventory level is shown in Fig. 3. The intervals with $x(t) > 0$ are those in which it is necessary to produce parts in advance, in order to satisfy future external demands whose rates are greater than the maximum speed U .

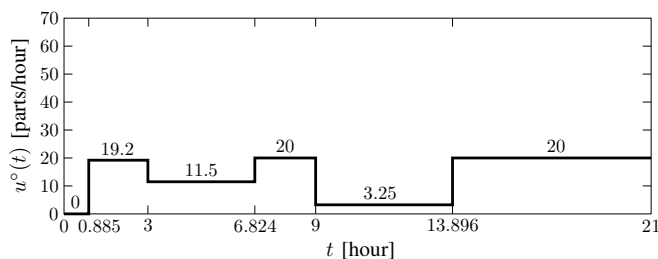


Fig. 2. Example – Optimal production rate $u^o(t)$.

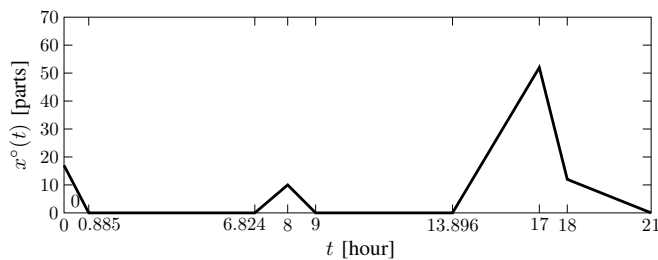


Fig. 3. Example – Optimal pattern of inventory level $x^o(t)$.

III. PERTURBATIONS ON THE SYSTEM STATE

The optimal open-loop solution provided in the previous section is determined on the basis of data that are known

at the beginning (for the entire planning horizon). However, for an on-line application, a close-loop solution would be surely preferable. As known, during the application of the optimal plan, it can happen that some perturbations occur to the system, making the off-line plan no more optimal or even no more feasible. For these reasons, it would be useful to determine optimal control strategies as functions of the current system state, that could be applied in presence of system behaviors affected by unmodelled perturbations to the system dynamics or to the model parameters.

First of all, it is necessary to make precise assumptions as regards the considered types of perturbations and the instants at which the system state is observed and the strategy is applied. The inventory level (system state) is supposed to be periodically measured and, after these measures, it can happen that the inventory level is different from the expected one. In particular, two cases must be distinguished:

- the inventory level is lower than the expected one (*negative perturbation*); this can be due to a malfunction in the production process (for which the machines have worked at a lower production effort or some finite products have been discarded, being faulty) or to an external demand that has resulted to be greater than the estimated one and then a higher quantity of finite products have been withdrawn from the inventory;
- the inventory level is greater than the expected one (*positive perturbation*); this can be due to a wrong setting of the production effort of the machines (for which the machines have worked at a higher production rate) or to an external demand lower than the estimated one, causing a surplus of finite products in the inventory.

The determination of the closed-loop strategies must be made considering the same cost terms adopted for the planning activity, i.e. setup, production, and holding costs, but in this case it is also necessary to introduce another cost term related to the case in which the demand is not completely satisfied, because the system perturbations (in particular, the negative perturbations) can provoke a shortage period, in which the external demand is not completely satisfied. The shortage cost is then defined as

$$C_{H^-} = \varphi^- \int_0^T \max\{0, -x(t)\} dt \quad (25)$$

being φ^- the unitary shortage cost per time unit. It is assumed φ^- sufficiently large so that the system is forced, even in the perturbed case, to satisfy the external demand.

The optimal control strategies which minimize setup, production, holding and shortage costs can be defined with reference to the *optimal path* of the inventory level, which is obtained by applying the optimal production speeds determined by solving Problem 1.

More specifically, when a negative perturbation occurs (i.e., the inventory level is lower than expected), the optimal control strategy is the one which allows the system state to return to the optimal path as soon as possible. This can be easily proven by considering that, in this case, the shortage

cost is surely paid (immediately, if the inventory level when the perturbation occurs is negative, or later, if the level is positive). Then, since the shortage cost has been supposed dominant with respect to the other costs, it is necessary to minimize the area provided by the integral in (25); this can be done by reaching the optimal path as soon as possible, taking into account the physical constraints of the system (first of all, the maximum production effort).

Also when a positive perturbation occurs, the optimal control strategy is the one which allows the system state returning to the optimal path as soon as possible, in order to minimize the extra-holding cost due to the unexpected higher number of parts in the inventory. However, in this case, it is necessary to take into account that a setup cost is paid if the machine is switched off in order to reach the optimal path as soon as possible. Then, two different strategies can be evaluated: either switching the machine off or continuing the production at a lower effort; the latter case is characterized by a higher extra-holding cost with respect to the extra-holding cost that is paid in the former case, but also the setup cost is paid in the former case. The best between the two strategies depends on the magnitude of the perturbation (other than, obviously, on the values of the parameters which characterize the system). It is worth noting that the latter is always advantageous if no technological constraint exists about the lowest speed that the production resource can keep without problems.

All these considerations are formalized in the next section where the closed-loop strategies, in case of negative and positive perturbations, are defined.

IV. OPTIMAL CONTROL STRATEGIES

Let $u^\circ(t)$ and $x^\circ(t)$, $0 \leq t \leq T$, be, respectively, the *optimal path* of the production speed (piece-wise constant function of t) and the *optimal path* of the inventory level (piece-wise linear function of t), which are obtained by using the results provided by Theorems 1 and 2. In particular, $x^\circ(t)$ represents the actual level of the inventory when no perturbation affects the system. However, when some perturbations occur, the actual level of the inventory may differ from $x^\circ(t)$. In this connection, consider a generic time instant $t^* \in [0, T)$ at which the level of inventory is measured, and let x^* be the measurement. Moreover, assume that a perturbation has occurred somewhere in $[0, t^*)$, such that it results $x^* \neq x^\circ(t^*)$; in other words, the parts contained in the inventory are less or more than expected.

The objective of this section is to determine, when a perturbation occurs, a piece-wise constant function of t representing the new optimal path of the production speed, namely $u^p(t, t^*, x^*)$, $t > t^*$, which minimizes the costs of the system. Function $u^p(t, t^*, x^*)$ is a closed-loop strategy, which is function of the actual system state at t^* , namely x^* . The function

$$x^p(t, t^*, x^*) = x^* + \int_{t^*}^t u^p(\eta, t^*, x^*) d\eta - \int_{t^*}^t e(\eta) d\eta \quad (26)$$

is the *optimal perturbed path* of the inventory level (piece-wise linear function of t), from t^* on.

A. Negative perturbation

In case the measurement x^* at t^* is such that it results $x^* < x^\circ(t^*)$, then the parts contained in the inventory are less than expected. The closed-loop strategy is the following:

- if $\phi(t^*, x^*) < T$:

$$u^p(t, t^*, x^*) = \begin{cases} U & t^* < t < \phi(t^*, x^*) \\ u^\circ(t) & \phi(t^*, x^*) \leq t \leq T \end{cases} \quad (27)$$

- if $\phi(t^*, x^*) = T$:

$$u^p(t, t^*, x^*) = U \quad t^* < t \leq T \quad (28)$$

where time instant $\phi(t^*, x^*)$ can be obtained by executing the following algorithm:

- 1: $d = x^\circ(t^*) - x^*$
- 2: $h \in \{1, \dots, P\} : \tau_{h-1}^\circ \leq t^* < \tau_h^\circ$
- 3: $h^s = h$
- 4: $\phi = t^*$
- 5: **while** $d > 0$ **and** $h \leq P$ **do**
- 6: $\delta = d - (U - u_{h^s}^\circ)(\tau_{h^s}^\circ - \phi)$
- 7: **if** $\delta \leq 0$ **then**
- 8: $\phi = \phi + \frac{d}{U - u_{h^s}^\circ}$
- 9: **else**
- 10: $\phi = \tau_{h^s}^\circ$
- 11: $h = h + 1$
- 12: **end if**
- 13: $d = \delta$
- 14: **end while**
- 15: $\phi(t^*, x^*) = \phi$
- 16: $h^e = h$

$\phi(t^*, x^*)$ is the time instant at which the strategy brings the perturbed path back to the optimal path, that is, the time instant at which the effects of the perturbation come to an end. It is apparent that, from $\phi(t^*, x^*)$ on, the system can continue with $u^\circ(t)$.

The extra-cost paid in consequence of such a type of perturbation is the difference between a negative-inventory cost and a positive-inventory gain (if any), that is

$$C_{\text{ex}} = \varphi^- \int_{t^*}^{\phi(t^*, x^*)} \max\{0, -x^p(t, t^*, x^*)\} dt - \varphi \int_{t^*}^{\phi(t^*, x^*)} [x^\circ(t) - \max\{0, x^p(t, t^*, x^*)\}] dt \quad (29)$$

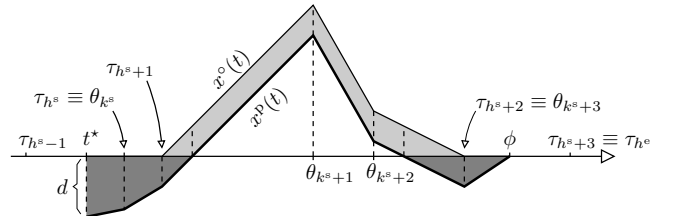


Fig. 4. Example of negative perturbation.

An analytical expression of such an extra-cost can be always determined on the basis of geometrical considerations.

As an example, with reference to the negative perturbation $d = x^\circ(t^*) - x^*$ illustrated in Fig. 4, the extra cost is

$$\begin{aligned}
C_{\text{ex}} &= \frac{1}{2} \varphi^-(d + \delta_{h^s})(\tau_{h^s} - t^*) \\
&+ \frac{1}{2} \varphi^-(\delta_{h^s} + \delta_{h^s+1})(\tau_{h^s+1} - \tau_{h^s}) + \frac{1}{2} \varphi^- \frac{\delta_{h^s+1}^2}{U - e_{k^s+1}} \\
&+ \frac{1}{2} \varphi^- \frac{\delta_{h^s+2}^2}{U - e_{k^s+3}} + \frac{1}{2} \varphi^- \frac{\delta_{h^s+2}^2}{U - u_{h^s+3}} \\
&- \varphi(\tau_{h^s+2} - \tau_{h^s+1}) + \frac{1}{2} \varphi \frac{\delta_{h^s+1}^2}{U - e_{k^s+1}} \\
&+ \frac{1}{2} \varphi \frac{\delta_{h^s+2}^2}{U - e_{k^s+3}}
\end{aligned} \tag{30}$$

where $\delta_{h^s} = d - (U - u_{h^s})(\tau_{h^s} - t^*)$, $\delta_{h^s+1} = \delta_{h^s} - (U - u_{h^s+1})(\tau_{h^s+1} - \tau_{h^s})$, and $\delta_{h^s+2} = \delta_{h^s+1}$ (the negative-inventory cost is proportional to the dark-gray area in Fig. 4, whereas the positive-inventory gain is proportional to the light-gray area).

The analytical expression of the extra-cost for the general case can be determined as well; however, it requires a “notational effort” which is out of the scope of the present paper. In any case, note that, in order to determine the closed-loop strategy it is only necessary to know the optimal paths $u^\circ(t)$ and $x^\circ(t)$, $t^* \leq t \leq T$, whereas to compute the extra-cost also the external demand e_k , $k = k^s, \dots, Q$, with k^s such that $\theta_{k^s-1} \leq t^* < \theta_{k^s}$, has to be taken into account.

B. Positive perturbation

In case the measurement x^* at t^* is such that it results $x^* > x^\circ(t^*)$, then the parts contained in the inventory are more than expected. In this case, the closed-loop strategy is the following:

- if $\phi(t^*, x^*) < T$ and $\varphi A \geq \sigma$:

$$u^p(t, t^*, x^*) = \begin{cases} 0 & t^* < t < \phi(t^*, x^*) \\ u^\circ(t) & \phi(t^*, x^*) \leq t \leq T \end{cases} \tag{31}$$

- if $\phi(t^*, x^*) < T$ and $\phi_\epsilon(t^*, x^*) < T$ and $\varphi A < \sigma$:

$$u^p(t, t^*, x^*) = \begin{cases} \epsilon & t^* < t < \phi_\epsilon(t^*, x^*) \\ u^\circ(t) & \phi_\epsilon(t^*, x^*) \leq t \leq T \end{cases} \tag{32}$$

- if $\phi(t^*, x^*) < T$ and $\phi_\epsilon(t^*, x^*) = T$ and $\varphi A < \sigma$:

$$u^p(t, t^*, x^*) = \epsilon \quad t^* < t \leq T \tag{33}$$

- if $\phi(t^*, x^*) = T$:

$$u^p(t, t^*, x^*) = 0 \quad t^* < t \leq T \tag{34}$$

where time instant $\phi(t^*, x^*)$ (resp., $\phi_\epsilon(t^*, x^*)$) can be obtained by executing the following algorithm:

- 1: $d = x^* - x^\circ(t^*)$
- 2: $h \in \{1, \dots, P\} : \tau_{h-1}^\circ \leq t^* < \tau_h^\circ$
- 3: $h^s = h$
- 4: $\phi = t^*$
- 5: **while** $d > 0$ **and** $h \leq P$ **do**
- 6: $\delta = d - u_h^\circ(\tau_h^\circ - \phi)$ (resp., $\delta = d - (u_h^\circ - \epsilon)(\tau_h^\circ - \phi)$)
- 7: **if** $\delta \leq 0$ **then**

- 8: $\phi = \phi + \frac{d}{u_h^\circ}$ (resp., $\phi = \phi + \frac{d}{u_h^\circ - \epsilon}$)
- 9: **else**
- 10: $\phi = \tau_h^\circ$
- 11: $h = h + 1$
- 12: **end if**
- 13: $d = \delta$
- 14: **end while**
- 15: $\phi(t^*, x^*) = \phi$ (resp., $\phi_\epsilon(t^*, x^*) = \phi$)
- 16: $h^e = h$ (resp., $h_\epsilon^e = h$)

and A is the difference between the additional inventory that is incurred by the system when producing at speed ϵ from t^* to $\phi_\epsilon(t^*, x^*)$ and the additional inventory which is incurred by the system when stopping the production from t^* to $\phi(t^*, x^*)$. A can be determined on the basis of geometric considerations. In the most general case, when $h^e > h^s + 1$, $h^e > h_\epsilon^e + 1$, and $h^e \leq P$, its value is

$$\begin{aligned}
A &= \frac{a^2}{2\epsilon} + \frac{ab}{u_{h^e}} + \frac{b^2 \epsilon}{2u_{h^e}^2} + \\
&+ \frac{1}{2} \left(a + \frac{b\epsilon}{u_{h^e}} + c_{u_{h^e}} \right) \left(\tau_{h^e} - \tau_{h^e-1} - \frac{b}{u_{h^e}} \right) \\
&+ \sum_{i=h^e+1}^{h_\epsilon^e-1} \frac{1}{2} (c_i - c_{i-1}) (\tau_i - \tau_{i-1}) + \frac{c_{h_\epsilon^e}^2}{2(u_{h_\epsilon^e} - \epsilon)}
\end{aligned} \tag{35}$$

where a , b , and c_j , $j = h^e, \dots, h_\epsilon^e - 1$, take the following values

$$a = \epsilon (\tau_{h^e-1} - t^*) \tag{36}$$

$$b = d - u_{h^s}(\tau_{h^s} - t^*) - \sum_{i=h^s+1}^{h^e-1} u_i(\tau_i - \tau_{i-1}) \tag{37}$$

$$c_{h^e} = \delta - (u_{h^e} - \epsilon) \left(\tau_{h^e} - \tau_{h^e-1} - \frac{b}{u_{h^e}} \right) \tag{38}$$

$$\begin{aligned}
c_j &= \delta - (u_{h^e} - \epsilon) \left(\tau_{h^e} - \tau_{h^e-1} - \frac{b}{u_{h^e}} \right) \\
&- \sum_{i=h^s+1}^{h_\epsilon^e-1} (u_i - \epsilon) (\tau_i - \tau_{i-1}) \quad j = h^e + 1, \dots, h_\epsilon^e - 1
\end{aligned} \tag{39}$$

with $d = x^* - x^\circ(t^*)$ and $\delta = a + \frac{b\epsilon}{u_{h^e}}$. Instead, in the case $h^e = h_\epsilon^e$, $h^e \leq P$, the value of A is

$$A = \frac{1}{2} a^2 \frac{u_{h^e}}{\epsilon(u_{h^e} - \epsilon)} + a b \frac{1}{u_{h^e} - \epsilon} + \frac{1}{2} b^2 \frac{\epsilon}{u_{h^e}(u_{h^e} - \epsilon)} \tag{40}$$

Finally, in the case $h^s = h^e = h_\epsilon^e$, $h^e \leq P$, the value of A is simply

$$A = \frac{1}{2} d^2 \frac{\epsilon}{u_{h^s}(u_{h^s} - \epsilon)} \tag{41}$$

(such three considered cases are not exhaustive; however, the other existing intermediate cases can be derived from the most general case).

If $\phi(t^*, x^*) < T$ and $\varphi A \geq \sigma$, the extra cost paid in consequence of such a type of perturbation is a positive-inventory cost plus a setup cost; otherwise, the extra cost is only the positive-inventory cost. Thus

- if $\phi(t^*, x^*) < T$ and $\varphi A \geq \sigma$:

$$C_{\text{ex}} = \sigma + \varphi \int_{t^*}^{\phi(t^*, x^*)} [x^{\text{P}}(t, t^*, x^*) - x^{\circ}(t)] dt \quad (42)$$

- if $(\phi(t^*, x^*) < T$ and $\varphi A < \sigma$) or $\phi(t^*, x^*) = T$:

$$C_{\text{ex}} = \varphi \int_{t^*}^{\phi_{\epsilon}(t^*, x^*)} [x^{\text{P}}(t, t^*, x^*) - x^{\circ}(t)] dt \quad (43)$$

(note that, when $\phi(t^*, x^*) = T$, also $\phi_{\epsilon}(t^*, x^*) = T$)

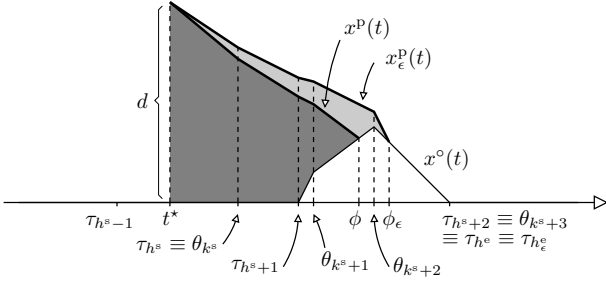


Fig. 5. Example of positive perturbation.

As in the case of negative perturbations, also in this case it is always possible to determine an analytical expression of the extra-cost, on the basis of geometrical considerations. As an example, with reference to the positive perturbation $d = x^* - x^{\circ}(t^*)$ illustrated in Fig. 5, the extra cost is:

- when stopping the production from t^* to $\phi(t^*, x^*)$:

$$\begin{aligned} C_{\text{ex}} &= \sigma + \frac{1}{2} \varphi (d + \delta_{h^s}) (\tau_{h^s} - t^*) \\ &+ \frac{1}{2} \varphi (\delta_{h^s} + \delta_{h^{s+1}}) (\tau_{h^{s+1}} - \tau_{h^s}) \\ &+ \frac{1}{2} \varphi (\delta_{h^{s+1}} + \lambda_{k^s+1}) (\theta_{k^s+1} - \tau_{h^{s+1}}) + \frac{1}{2} \varphi \frac{\lambda_{k^s+1}^2}{U} \end{aligned} \quad (44)$$

where $\delta_{h^s} = d - u_{h^s}(\tau_{h^s} - t^*)$, $\delta_{h^{s+1}} = \delta_{h^s} - u_{h^{s+1}}(\tau_{h^{s+1}} - \tau_{h^s})$, and $\lambda_{k^s+1} = \delta_{h^{s+1}} - U(\theta_{k^s+1} - \tau_{h^{s+1}})$ (the extra positive-inventory cost is proportional to the dark-gray area in Fig. 5);

- when producing at speed ϵ from t^* to $\phi_{\epsilon}(t^*, x^*)$:

$$\begin{aligned} C_{\text{ex}} &= \frac{1}{2} \varphi (d + \delta_{h^s}) (\tau_{h^s} - t^*) \\ &+ \frac{1}{2} \varphi (\delta_{h^s} + \delta_{h^{s+1}}) (\tau_{h^{s+1}} - \tau_{h^s}) \\ &+ \frac{1}{2} \varphi (\delta_{h^{s+1}} + \lambda_{k^s+1}) (\theta_{k^s+1} - \tau_{h^{s+1}}) \\ &+ \frac{1}{2} \varphi (\lambda_{h^s+1} + \lambda_{k^s+2}) (\theta_{k^s+2} - \theta_{h^s+1}) + \frac{1}{2} \varphi \frac{\lambda_{k^s+2}^2}{U} \end{aligned} \quad (45)$$

where in this case $\delta_{h^s} = d - (u_{h^s} - \epsilon)(\tau_{h^s} - t^*)$, $\delta_{h^{s+1}} = \delta_{h^s} - (u_{h^{s+1}} - \epsilon)(\tau_{h^{s+1}} - \tau_{h^s})$, $\lambda_{k^s+1} = \delta_{h^{s+1}} - (U - \epsilon)(\theta_{k^s+1} - \tau_{h^{s+1}})$, and $\lambda_{k^s+2} = \lambda_{k^s+1} - (U - \epsilon)(\theta_{k^s+2} - \theta_{k^s+1})$ (the extra positive-inventory cost is proportional to the sum of the dark-gray area and the light-gray area in Fig. 5).

Then, in the example of positive perturbation illustrated in Fig. 5, the machine is stopped if the light-gray area weighted by the unitary positive-inventory cost φ is greater than or equal to the setup cost σ . Since in this example $h^e = h^e_{\epsilon} = h^s + 2$, then, in accordance with (40), the machine is stopped when

$$\begin{aligned} \frac{1}{2} \varphi a^2 \frac{u_{h^s+2}}{\epsilon(u_{h^s+2} - \epsilon)} + \varphi a b \frac{1}{u_{h^s+2} - \epsilon} \\ + \frac{1}{2} \varphi b^2 \frac{\epsilon}{u_{h^s+2}(u_{h^s+2} - \epsilon)} \geq \sigma \end{aligned} \quad (46)$$

in which $a = \epsilon(\tau_{h^s+1} - t^*)$ and $b = d - u_{h^s}(\tau_{h^s} - t^*) - u_{h^s+1}(\tau_{h^s+1} - \tau_{h^s})$.

Remark 1: In writing the closed-loop strategies (31) and (32), it has been implicitly assumed that there is a technological constraint on the lowest speed that the production resource can keep without problems (e.g., without having unexpected shutdowns), and that such speed is equal to ϵ ; in other words, with respect to the generic interval $[\tau_{h-1}, \tau_h)$, it is either $u_h = 0$ or $\epsilon \leq u_h \leq U$. Without such technological constraint, it can be shown that a production speed sufficiently low so that $\varphi A < \sigma$ always exists (as a matter of fact, $A \rightarrow 0$ when the production rate tends to 0), and then it is always convenient to not stop the resource and avoid the payment of the setup cost. However, in order to exploit the results in [17], it is implicitly assumed that also the positive external demand is greater than or equal to ϵ .

C. Example – Part II

Consider again the inventory-production system illustrated Subsection II-D. Assume that the inventory is measured at $t^* = 5.8$ hours and that the level measures $x^* = -10$. A negative perturbation (which has occurred somewhere in the interval $[0, 5.8)$) is detected and then an optimal control strategy can be defined in order to minimize the extra-cost due to the perturbation. By using the results included in Subsection IV-A, it turns out $\phi(5.8, -10) = 9.077$; then

$$u^{\text{P}}(t, 5.8, -10) = \begin{cases} 20 & t^* < t < 9.077 \\ u^{\circ}(t) & 9.077 \leq t \leq 21 \end{cases} \quad (47)$$

and the new optimal path of the production speed is illustrated in Fig. 6. In Fig. 7, the optimal perturbed path of the inventory level, in the interval $[5, 10]$, which is obtained when using $u^{\text{P}}(t, 5.8, -10)$, is reported.

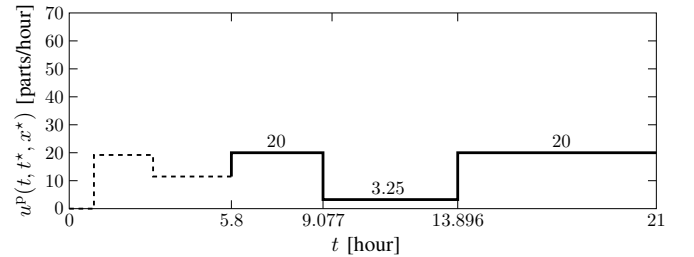


Fig. 6. New optimal path of the production speed, in case of a negative perturbation at $t^* = 5.8$, at which $x^* = -10 < x^{\circ}(t^*) = 0$.

Assume now that a different perturbation has occurred, so that it results $x^* = 10$ at $t^* = 5.8$ (positive perturbation). In

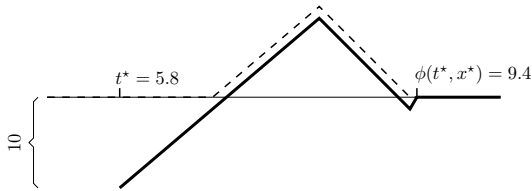


Fig. 7. Perturbed (solid line) and optimal (dashed line) paths of the inventory level, in case of a negative perturbation at $t^* = 5.8$, at which $x^* = -10 < x^o(t^*) = 0$ (only the interval $[5, 10]$ is depicted).

this case, in order to define the optimal control strategy, it is necessary to take into consideration the setup cost σ , the unitary holding cost φ , and the lowest speed ϵ ; such values are assumed to be equal to 1 (euros), 2 (euros/part/hour), and 2 (parts/hour), respectively. In this case, by using the results included in Subsection IV-B, it turns out $A = 0.914$ and then $\varphi A > \sigma$. Since, $\phi(5.8, 10) = 6.670$, the optimal control strategy is

$$u^p(t, 5.8, 10) = \begin{cases} 0 & t^* < t < 6.670 \\ u^o(t) & 6.670 \leq t \leq 21 \end{cases} \quad (48)$$

and the new optimal path of the production speed is illustrated in Fig. 8. In Fig. 9, the optimal perturbed path of the inventory level, in the interval $[5, 8]$, which is obtained when using $u^p(t, 5.8, 10)$, is reported.

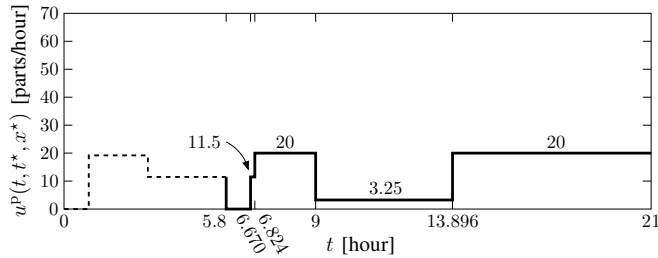


Fig. 8. New optimal path of the production speed, in case of a positive perturbation at $t^* = 5.8$, at which $x^* = 10 > x^o(t^*) = 0$.

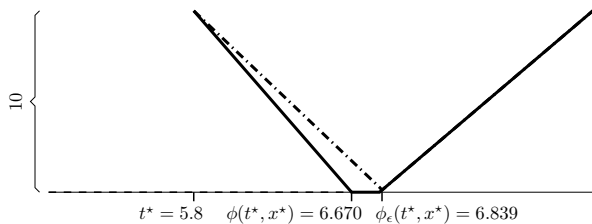


Fig. 9. Perturbed (solid and dot-dashed lines) and optimal (dashed line) paths of the inventory level, in case of a positive perturbation at $t^* = 5.8$, at which $x^* = 10 > x^o(t^*) = 0$ (only the interval $[5, 8]$ is depicted). The solid line refers to the case in which the machine is switched off, whereas the dot-dashed line to the case with production speed set to $\epsilon = 2$ parts/hour.

V. CONCLUSIONS

In this paper a class of inventory-production systems is considered, with the objective of finding optimal control strategies to be applied when the system is affected by perturbations on the inventory level. In particular, the considered model represents a manufacturing process in which the external demand has a piecewise constant pattern changing rate at asynchronous time instants; the state variable is the

inventory level and the control variable is the production rate, again with a piecewise constant path. In the paper, optimal (closed-loop) control strategies are found with reference to an optimal path (found by solving a parametric, nonlinear, combinatorial optimization problem), in order to minimize the extra-cost due to the system perturbations.

It is worth finally noting that the methodology proposed in this paper to find the optimal control strategies can be also applied to different models of inventory-production systems, in which the inventory level grows and decreases linearly with time, such as the models characterized by a non-negative external demand and/or by the presence of a constraint on the minimum production speed. The determination of the optimal open-loop solution for such different models are currently under investigation.

REFERENCES

- [1] B. Karimi, S. Fatemi Ghomi, and J. Wilson, "The capacitated lot sizing problem: a review of models and algorithms," *Omega*, vol. 31, no. 5, pp. 365–378, 2003.
- [2] N. Brahimi, S. Dauzere-Peres, N. Najid, and A. Nordli, "Single item lot sizing problems," *European Journal of Operational Research*, vol. 168, no. 1, pp. 1–16, 2006.
- [3] H. Wagner and T. Whitin, "Dynamic version of the economic lot size model," *Management Science*, vol. 5, no. 1, pp. 89–96, 1958.
- [4] M. Ortega and L. Lin, "Control theory applications to the production-inventory problem: a review," *International Journal of Production Research*, vol. 42, no. 11, pp. 2303–2322, 2004.
- [5] S. Benjaafar, M. ElHafsi, and T. Huang, "Optimal control of a production-inventory system with both backorders and lost sales," *Naval Research Logistics*, vol. 57, no. 3, pp. 252–265, 2010.
- [6] E. Khmelnitsky and Y. Gerchak, "Optimal control approach to production systems with inventory-level-dependent demand," *IEEE Transactions on Automatic Control*, vol. 47, no. 2, pp. 289–292, 2002.
- [7] F. Chu and C. Chu, "Single-item dynamic lot-sizing models with bounded inventory and outsourcing," *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 38, no. 1, pp. 70–77, 2008.
- [8] W. Jaruphongsa and C.-Y. Lee, "Dynamic lot-sizing problem with demand time windows and container-based transportation cost," *Optimization Letters*, vol. 2, no. 1, 2008.
- [9] C. Bannister and R. Kaye, "A rapid method for optimization of linear systems with storage," *Operations Research*, vol. 39, no. 2, pp. 220–232, 1991.
- [10] R. Akella and P. Kumar, "Optimal control of production rate in a failure prone manufacturing system," *IEEE Transactions on Automatic Control*, vol. 31, no. 1, pp. 116–126, 1986.
- [11] S. Gershwin, *Manufacturing Systems Engineering*, Prentice-Hall, Ed., 1994.
- [12] F. Martinelli and P. Valigi, "Hedging point policies remain optimal under limited backlog and inventory space," *IEEE Transactions on Automatic Control*, vol. 49, no. 10, pp. 1863–1869, 2004.
- [13] M. Gazarik and Y. Wardi, "Optimal release times in a single server: an optimal control perspective," *IEEE Transactions on Automatic Control*, vol. 43, no. 7, pp. 998–1002, 1998.
- [14] D. L. Pepyne and C. G. Cassandras, "Modeling, analysis, and optimal control of a class of hybrid systems," *Discrete Event Dynamic Systems*, vol. 8, no. 2, pp. 175–201, 1998.
- [15] J. Moon and Y. Wardi, "Optimal control of processing times in single-stage discrete event dynamic systems with blocking," *IEEE Transactions on Automatic Control*, vol. 50, no. 6, pp. 880–884, 2005.
- [16] D. Giglio, R. Minciardi, S. Sacone, and S. Siri, "Optimal control of production processes with variable execution times," *Discrete Event Dynamic Systems: Theory and Applications*, vol. 19, no. 3, pp. 423–448, 2009.
- [17] D. Giglio, S. Sacone, and S. Siri, "Asynchronous regulation of service speed in inventory-production systems with time-varying positive demand," in *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, 2011, pp. 5874–5880.