Feedback control for priority rules in re-entrant semiconductor manufacturing

Fenglan He a, Dieter Armbruster b,⇑, Michael Herty c, Ming Dong d

a Innovative Ocean and Energy System Group, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, PR China
b School of Mathematical and Statistical Sciences, Arizona State University, Tempe, USA
c RWTH Aachen University, Aachen, Germany
d Antai College of Economics & Management, Shanghai Jiao Tong University, PR China

A R T I C L E   I N F O
Article history:
Received 10 September 2014
Received in revised form 6 March 2015
Accepted 20 March 2015
Available online 11 May 2015

Keywords:
Re-entrant production
Feedback control
Partial differential equations
Dispatch policy control

A B S T R A C T
Typical semiconductor production is re-entrant and hence requires priority decisions when parts compete for production capacity at the same machine. A standard way to run such a factory is to start to plan and to finish according to demand. Often this results in a push policy where early production steps have priority over later production steps at the beginning of the production line and a pull policy where later steps have priority at the end of the production line. The point where the policies switch is called the push–pull point (PPP). We develop a control scheme based on moving the PPP in a continuum model of the production flow. We show that this control scheme significantly reduces the mismatch between demand and production output. The success of the control scheme as a function of the frequency of control action is analyzed and optimal times between control actions are determined.

1. Introduction

Semiconductor manufacturing is arguably one of the most advanced industrial production processes. Due to its long cycle times (order of several weeks) and its re-entrant flow topology it poses many challenges for production planning and control. Especially for commodity chip production when only one type of chip is produced and the product mix cannot be altered, there are not a lot of ways to influence the production flow beyond the starts policy.

However, re-entrant production where a machine is used to serve several steps in the production schedule of a chip, allows the use of a dispatch policy to set priorities for all the steps that need to be performed at that machine. A common dispatch policy in this case is the push policy, also known as first buffer first served policy, which gives priority to earlier production steps over later production steps. A pull policy, also known as shortest-expected-remaining-process-time policy, gives priority to later production steps over earlier production steps. Push policies are usually used at the front of a production line, whereas pull policies are used at the back of the factory. The step where push policy switches to pull policy is called the push–pull point (PPP).

In a discrete simulation (DES) model of a semiconductor factory Perdaen et al. [1] study the response of a production line to demand fluctuations that occur on a much shorter timescale than the cycle time of the factory. They show that the...
mismatch between a fluctuating demand and the factory production can be reduced dramatically relative to many other static production scenarios when they couple a heuristics for moving the PPP with a CONWIP starts policy [2].

While DES can be made extremely accurate and their use in practice is widespread, they are extremely time consuming, both to run and to maintain. In addition they have the usual issues of simulation studies: It is difficult to extrapolate their generality from the specific parameters that are used and insights into the simulation structure and its dependence on parameters is hard to get. Therefore, in recent years aggregate production models have been developed. Initially they were based on a continuum of product leading to queuing network models [3] which led to the concept of the clearing function, relating output in a period to workload during the period (see [4] for a state of the art review). More recently models based on transport equations considering production flow as a fluid flow in a continuum in product as well as in production steps have been discussed. The latter approach leads to partial differential equations (PDEs) describing the flow of the product density and the associated flow velocity [5].

In [6] a PDE model incorporating a PPP based dispatch policy was presented. The current paper improves this model and develops a linearized control algorithm. The current model is a feedback control model, tracking the outflux to a given known demand sequence. We will analyze the PDE model with this linearized control and compare it to fixed PPP schemes and heuristics similar to the one used in [1] that are based on giving high priority to the lots needed to satisfy the demand in the next time unit.

The general topic of production planning has a long history and the literature is large and diverse. For the more specific topic of production planning in semiconductor factories there is a comprehensive recent monograph by Mönch et al. [7] which presents an overview over real world approaches to dispatching rules, order releases and production planning at different scales from short term order releases to enterprise wide planning. A complete and exhaustive definition of production planning in its full generality is discussed by Kempf et al. [8]. Armbruster and Uzsoy’s review [9] is closest to our approach focussing on the interplay between discrete event simulations and aggregate models based on clearing functions and partial differential equations. Mathematical approaches to control at the aggregate level are discussed in [10,11]. All the previous approaches dealing with aggregate models were typically controlling the order release or starts policies. To our knowledge, none considered integrating the dispatch policies into an aggregate model. An optimization approach, corresponding to an open loop control problem for production planning in aggregate models integrating release and dispatch policies has been discussed in [12].

2. PDE model

As discussed in [5] and recently summarized in [9] we consider a work-in-progress (wip) density \( \rho(x, t) \) where \( x \) describes the continuum of production stages for a product. Hence \( \rho(0, t) \) is the wip density at the beginning production step and \( \rho(1, t) \) the wip density at the final production step. Mass conservation (known for a queuing system as Little’s law) is given by the continuity equation

\[
\frac{\partial \rho}{\partial t}(x, t) + \frac{\partial F}{\partial x}(x, t) = 0, \tag{1}
\]

where \( F(x, t) = \rho(x, t) v(x, t) \) is the associated flux of the production density and \( v(x, t) \) is the velocity of a part at stage \( x \) moving through the factory at time \( t \). Note that \( F \) is also known as the production rate, its units are parts/time. All the modeling efforts have been targeted to determine the best suited functional form of this flux. In steady state, the flux becomes constant representing the output of the production unit. The steady state flux is known as the clearing function [3] representing the relationship between the expected output of a production resource in a given planning period as a function of the workload.

Karmarkar [3] proposed a nonlinear clearing function where output increases as a concave nondecreasing function of \( W \), representing the output of the production unit. The steady state flux is known as the clearing function [3] representing the relationship between the expected output of a production resource in a given planning period as a function of the workload.

In the case of the clearing function, \( F \) represents the mean production rate for a production line in equilibrium as a function of the mean wip. In the case of \( v(W) \), this represents the mean velocity \((\frac{\mu}{\text{mean cycle time}})\) as a function of the mean wip. Hence Eq. (1) is based on a quasi-steady state assumption restricting its validity to slowly varying influxes.

A typical velocity model that is based on treating the whole factory as a single M/M/1 queue [13] is given by

\[
v(W) = \frac{\mu}{1 + W}, \tag{3}
\]
where $\mu$ is the mean processing rate of the M/M/1 queue, representing in the continuum case the speed of a part moving through an empty factory without waiting. Hence the steady states for the PDE Eq. (1) and a flux (2) for and a time–independent influx $\lambda$ are

$$\rho = W, \quad \lambda = \frac{\mu W}{1 + W}, \quad W = \frac{\lambda}{\mu - \lambda}, \quad \nu = \mu - \lambda, \quad CT := \frac{1}{\nu} = \frac{1}{\mu - \lambda}. \quad (4)$$

These are all the familiar relations for the mean queue length, the arrival rate, the production rate and the cycle time for an equilibrium distribution of an M/M/1 queue.

### 2.1. Continuum model with push–pull dispatch policies

We define $\xi(t)$ to be the PPP – for $x < \xi$ the dispatch policy is a push policy for $x > \xi$ it is a pull policy. We simplify the impact of the dispatch policy by assuming that the velocity in each of the two regions is constant and modeled by the queuing model Eq. (4). Effectively this implies that $\xi$ separates the high priority region (downstream) from the low priority region (upstream) and within each region we have a random dispatch policy.

Thus we have the velocity in the pull region depending on the total wip in the pull region given as

$$\nu^{\text{pull}} = \frac{v_0}{1 + C^{\text{pull}} W^{\text{pull}}}, \quad W^{\text{pull}} = \int_{\xi}^{1} \rho(x, t) \, dx. \quad (5)$$

At the same time, since stages in the pull region have priority over stages in the push region, all the wip in the factory competes for production capacity. Hence the velocity in the push region is given as

$$\nu^{\text{push}} = \frac{v_0}{1 + C^{\text{push}} W}, \quad W = \int_{0}^{1} \rho(x, t) \, dx. \quad (6)$$

Here $C^{\text{pull}}$ and $C^{\text{push}}$ are possibly factory dependent constants with $C^{\text{pull}} < C^{\text{push}}$, reflecting the fact that a factory that is run completely in pull mode (i.e. $\xi = 0$) has a higher throughput for the same wip than a factory run completely in push mode (i.e. $\xi = 1$) (see c.f. [1]). We define $C^{\text{push}} = C$ and $C^{\text{pull}} = 1$ for simplicity. Hence $C^{\text{push}} > 1$ represents the relative impact that a part has on the total flux in push and pull policies, respectively. The flux in pull region is given by $F = \rho^{\text{pull}} \nu^{\text{pull}}$ and similarly for the push region. In steady state the flux is $\lambda$ everywhere and the two constant density levels are called $\rho^{\text{pull}}$ and $\rho^{\text{push}}$. Accordingly the corresponding wips are $W^{\text{pull}} = \rho^{\text{pull}} (1 - \xi)$ and $W = \rho^{\text{push}} \xi + (1 - \xi) \rho^{\text{pull}}$. The wip densities are determined by solving the two equations

$$\frac{\lambda}{\rho^{\text{pull}}} = \frac{v_0}{1 + C^{\text{pull}} \xi (1 - \xi)}, \quad (7)$$

$$\frac{\lambda}{\rho^{\text{push}}} = \frac{v_0}{1 + C \rho^{\text{push}} \xi + \rho^{\text{pull}} (1 - \xi)}. \quad (8)$$

leading to a wip profile that depends on the position $\xi$ of the PPP:

$$\rho^{\text{pull}}_{ss}(x, t) = \frac{\lambda}{v_0 - \lambda (1 - \xi(t))}, \quad (9)$$

$$\rho^{\text{push}}_{ss}(x, t) = \frac{\lambda [v_0 + \lambda (1 - \xi(t)) (C - 1)]}{(v_0 - \lambda \xi(t) C) (v_0 - \lambda (1 - \xi(t)))}, \quad \text{and} \quad (10)$$

$$\rho_{ss}(x, t) = \rho^{\text{push}}_{ss}(x) (1 - H(x - \xi(t))) + \rho^{\text{pull}}_{ss}(x) H(x - \xi(t)), \quad (11)$$

with $H(x) = 1$ for $x > 0$ and $H(x) = 0$ otherwise, denoting the Heaviside function.
Fig. 1 shows the discontinuous steady states for different values of $\zeta$ and $\nu_0 = 1$, $C = 1.2$, $\lambda = 0.7$. Note that the density in the pull region is always higher than in the push region and that the discontinuity occurs at the position of the PPP.

This model has the following limiting behaviors: As $\zeta \to 0$ the system is run completely in pull dispatch mode and hence is characterized by

$$
\lim_{\zeta \to 0} \rho^{\text{pull}} = \frac{\lambda}{\nu_0 - \lambda},
\lim_{\zeta \to 0} v^{\text{pull}} = v_0 - \lambda,
\lim_{\zeta \to 0} q^{\text{push}} = \frac{\lambda [v_0 + \lambda (C - 1)]}{v_0 (v_0 - \lambda)},
\lim_{\zeta \to 0} v^{\text{push}} = \frac{v_0}{1 + C \rho^{\text{pull}}} = \frac{v_0 (v_0 - \lambda)}{v_0 + \lambda (C - 1)}.
$$

Similarly, as $\zeta \to 1$ the system is run completely in push dispatch mode and is characterized by

$$
\lim_{\zeta \to 1} \rho^{\text{pull}} = v_0,
\lim_{\zeta \to 1} v^{\text{pull}} = \frac{v_0}{\nu_0 - \lambda},
\lim_{\zeta \to 1} q^{\text{push}} = \frac{\lambda}{v_0 - \lambda C},
\lim_{\zeta \to 1} v^{\text{push}} = v_0 - \lambda C.
$$

In particular, since $C > 1$, the pull policy has a higher limiting flux than the push policy.

Finally, we consider a time-dependent PPP $\zeta = \zeta(t)$. Using the Eqs. (5) and (6) for the velocities we obtain

$$
\frac{\partial \rho}{\partial t} (x, t) + v^{\text{push}} \frac{\partial \rho}{\partial x} (x, t) = 0 \quad \text{for } x < \zeta,
\frac{\partial \rho}{\partial t} (x, t) + v^{\text{pull}} \frac{\partial \rho}{\partial x} (x, t) = 0 \quad \text{for } x > \zeta.
$$

The identity of the right and left hand limits at $\zeta(t) \pm$ of the flux

$$
\lim_{x \to \zeta^-} \rho(x, t) v^{\text{push}} (x, t) = \lim_{x \to \zeta^+} \rho(x, t) v^{\text{pull}} (x, t),
$$

guarantees mass conservation across the PPP.

2.1.1. Simulations

Fig. 2 shows simulation results of the PDE model Eq. (14) for a $\nu_0 = 8$ and a utilization of 75%: (a) shows three PPP jumps $\zeta = 0.5 \to 0.2 \to 0.7 \to 0.5$ at times $t = 1, 2, 3$; (b) shows the output and (c) the velocity as a function of time; (d) shows the wip profiles $\rho(x)$ just before the jumps at $t = 1, 2, 3$ and 4. Since $\nu_0$ is quite high, the perturbation generated by the jump in $\zeta$ moves out of the factory very quickly leading to spiking outputs and to wip profiles that have decayed to the steady state profiles by the time the PPP jumps again.
If we make the cycle time in the factory much longer, then the transition is not between steady states any more: Fig. 3 shows simulation results of the PDE model Eq. (14) for a $v_0 = 1$ and a utilization of 70% for the three jumps of the PPP shown in Fig. 2(a). Fig. 3(a) shows the output and (b) the velocity as a function of time. In this case the cycle time for a loaded factory is much larger than one time unit. Hence the transitions between PPP jumps are not from one steady state to another as can be seen from the wip profiles $q(x)$ just before the jumps at $t = 1, 2, 3$ in Fig. 3(c).

2.2. Feedback control

We are interested to change the position of the PPP to influence the outflux of the factory at discrete times corresponding to re-planning periods. We call the discrete time intervals $s_i$ and define a time sequence $t_i = i s, i = 0 \ldots N$. Then the production rate of the factory is given by $q(\xi(t), t) v_{pull}(t)$ where $\rho(x,t)$ (and subsequently $\rho(1,t)$ and $v(t)$) are found by successively solving the PDE (14) with $\rho(x,t_i)$ being the initial condition on the time interval $t \in [t_i, t_{i+1}]$ and $\rho(x,t_{i+1})$ the initial condition for the next time interval.

Let $D_i$ be the total demand in the $i$th time interval to be delivered at time $t_i$. The goal of the factory production manager is to deliver exactly $D_i$ units in the time interval between $t = t_{i-1}$ and $t_i$. We denote this output of the factory in the $i$th time interval with $O_i$. Using the PDE (14) the output can be written as

$$O_i = \int_{t_{i-1}}^{t_i} v_{pull}(t) \rho(1,t) dt.$$  \hspace{1cm} (16)

The objective is to find a policy for moving the PPP at time $\xi(t)$ given the state of the factory $\rho(x,t_{i-1})$ and the previous time and the demand in the current interval of interest $D_i$. This leads to a closed loop or feedback control question. In order to understand the impact of moving the PPP, we assume that at $t = t_{i-1}$, for $\xi(t_{i-1}) = \xi_{i-1}$ the factory is in steady state, given by a wip profile $\rho_{ss}(x,t_{i-1})$ as in Eq. (11). To derive a feedback law that determines the best new position of the PPP $\xi(t)$ we linearize the PDE (14) in the following way: We consider the linear transport equation

$$\frac{\partial \rho}{\partial t} (x,t) + b(t) \frac{\partial \rho}{\partial x} (x,t) = 0,$$  \hspace{1cm} (17)
\[
\rho(x, t_{i-1}) = \rho_{ss}(x, t_{i-1}) \quad \text{initial condition}
\]

with \(\bar{v}(t_i)\) a velocity that is constant over the time interval \([t_{i-1}, t_i]\) given by the pull velocity for the wip after the jump of the PPP from \(\zeta(t_{i-1})\) to an arbitrary state \(\bar{\zeta}\), i.e.

\[
\bar{v}(t_i) = \frac{v_0}{1 + C_{\text{pull}} W(t_i)},
\]

\[
W(t_i) = \int_{\bar{\zeta}(t_i)}^{\zeta(t_i)} \rho(x, t_{i-1})dx.
\]

The solution for the outflux of the linear PDE is simply given by moving the initial profile downstream

\[
F(t) = \begin{cases} 
\bar{v}(t_i)\rho(1, t) = \bar{v}(t_i)\rho_{ss}(1 - \bar{v}(t_i)t, t_{i-1}) & \text{for } 0 \leq t \leq \frac{1}{\bar{v}(t_i)}, \\
\lambda & \text{for } t > \frac{1}{\bar{v}(t_i)}. 
\end{cases}
\]

In order to find the correct new PPP position we therefore need so solve the equality

\[
D_i = \int_{t_{i-1}}^{t_i} F(s)ds,
\]

for the position \(\bar{\zeta}\) under the constraints \(0 \leq \bar{\zeta} \leq 1\). The new position of the PPP is then \(\zeta(t_i) = \bar{\zeta}\) and clearly this choice depends on \(D_i\) and \(\rho(x, t_{i-1})\) as well as the evolution of the equation. It is therefore a nonlinear feedback law. We summarize the feedback law in a Lemma.

**Lemma 2.1.** Given the demand \(D_i\) and the state of the factory \(\rho(x, t_{i-1})\) for \(0 \leq x \leq 1\). The feedback law for the movement of the PPP is given by
• If $\tilde{\xi} < 0$, then set $\tilde{\xi}(t_i) = 0$,
• If $0 < \tilde{\xi} < 1$, then set $\tilde{\xi}(t_i) = \tilde{\xi}$,
• If $\tilde{\xi} \geq 1$, then set $\tilde{\xi}(t_i) = 1$.

where $\tilde{\xi}$ is the solution to Eq. (22).

3. Analyzing the feedback law

3.1. Matching production to demand

The aim of the feedback law is to use the changes in priority rules associated with moving the PPP to generate an outflux over a time period $\tau$ that is as close as possible to the demand over that time period. Clearly changing the priority rules will not change the total amount of wip in the factory, hence the feedback control will not be able to track any long term drift in the demand that requires a change in the mean start rate into the factory. Hence our fundamental experimental setup is characterized by a demand sequence $D_i$, $i = 1 \ldots N$ generated by a random variable $d_i$ which is the demand rate in the time interval $t_i$ such that

$$D_i = d_i \tau.$$  \hspace{1cm} (23)

The variable $d_i$ is a random variable sampling from a uniform distribution $U(\bar{d} - \delta d, \bar{d} + \delta d)$ such that on average the demand rate is $\bar{d}$ and hence the start rate into the factory is $\lambda = \bar{d}$. For a predetermined sequence $\{D_i\}$ we then solve the PDE Eq. (14) with an initial condition given by the steady state wip distribution for the constant $\lambda$ and the feedback law given in Lemma 2.1 and compare the production sequence $\{O_i\}$ with the demand sequence.

To average over the sample sequences we create 100 such demand sequences $\{D_i\}_k$, $k = 1 \ldots 100$. Hence the demand in time interval $i$ for the sample $k$ is given by $D_i^k$ and the corresponding output of the factory is denoted by $O_i^k$. For each sequence we calculate the average error rate between demand and production over the time horizon $T = N\tau$

$$\bar{e}^k(\tau) = \frac{\sum_{i=1}^N |O_i^k - D_i^k|}{T}.$$  \hspace{1cm} (24)

The mean errors over all samples

$$\bar{e}(\tau) = \frac{\sum_{k=1}^{100} \bar{e}^k(\tau)}{100}.$$  \hspace{1cm} (25)

will be our main observable.

We are interested in two questions

1. Is the mean error rate $\bar{e}$ for the dispatch control via the linearized feedback (2.1) lower than for other dispatch rules and by how much?
2. How does the mean error rate $\bar{e}(\tau)$ depend on the time interval $\tau$?

Our feedback rule is based on the assumption that the system is in steady state within one re-planning interval $\tau$. Figs. 2 and 3 have shown that this depends crucially on $\nu_0$, which in turn determines the cycle time of the factory. Hence for a given $\nu_0$ we expect that for very short $\tau$, the feedback control will not perform well.

In all the following simulations and figures, we chose $\nu_0 = 8$, $\lambda = \bar{d}$ and $C = 1.2$. Fig. 4 shows a typical result for $\lambda = 6$ (giving a utilization of $u = 75\%$) and a coefficient of demand variations of $c_v(d) = 0.6$.

We analyze four control policies: A pure pull policy, i.e. the PPP stays constant at $\xi = 0$; a pure push policy, i.e. the PPP stays constant at $\xi = 1$; the linearized feedback control algorithm; and a policy (called PPP in the figure) that moves the PPP in such a way that the wip downstream from the new PPP $\xi_t$ is equal to the demand $D_t$ in the next time period.

We observe that the linearized feedback control is always better than all the other algorithms and that the second best algorithm is the pure pull algorithm. Also, as $\tau$ increases, all algorithms have the same error since for large $\tau$ the dispatch rules do not matter and the outflux becomes the mean outflux $O_i = \bar{d} \tau$ and the error $|O_i - D_i|$ is just given by the noise level of $D_i$. The control algorithms differ only in their impact on the outflux for small $\tau$. While the linearized feedback control improves the outflux for $0 < \tau < 0.5$, the maximal improvement over the pure pull algorithm is about 30% and shows up at around $\tau = 0.3$. Note that treating the whole factory as an $M/M/1$ queue with a production rate of $\mu = 0$ and an inflow of $\lambda = 6$ leads to a steady state cycle time (cf. Eq. (4)) of $CT = 0.5$, i.e. moving the PPP improves production planning if the re-planning period is less than the cycle time and it works best if it is about 1/2 of the cycle time of the factory.

There is one additional interesting feature in Fig. 4. As $\tau \to 0$, the linearized feedback control does not reduce the error to zero, as one could have expected from a linearization. The reason is that every time the PPP moves, it creates a new discontinuity in the wip profile. Hence the assumption of a steady state wip profile is not correct. For small re-planning time
intervals, the changes in the demand will necessarily be small and hence the error generated by the wrong wip profile will be of the order of the change in the demand, making feedback control impossible.

3.2. The influence of the utilization and the demand variability

Moving the PPP creates additional discontinuities and the linearization is an approximation to the truly nonlinear behavior. Hence the feedback control algorithm creates its own noise. Anytime the system noise is of the same order of magnitude, the feedback control will not work well but mostly chase the noise that is created by the algorithm itself. This is in particular true for small demand variability and low utilization.

However, as the demand variability grows the benefits of the feedback control become apparent: We measure the relative improvement \( e_r \), of the mean error of the feedback control vs. the second best algorithm, the pure pull policy

\[
e_r = \frac{\bar{e}_{\text{pull}} - \bar{e}_{\text{feedback}}}{\bar{e}_{\text{pull}}}. \tag{26}
\]

Fig. 5(a) plots \( e_r \) at the optimal re-planning time \( \tau \) for a fixed utilization of \( u = 75\% \) as a function of the coefficient of variation of the demand signal.

Since moving the PPP is essentially a reallocation of priority to move excess capacity in the factory to high wip regions the feedback control algorithm will show diminishing returns as the utilization increases. In particular, for very high utilization, there is little excess capacity that can be re-allocated and hence the improvement will become small. Fig. 5(b) plots \( e_r \) at the optimal re-planning time \( \tau \) for a fixed \( c_v(d) = 0.6 \) of the demand signal as a function of the utilization.

4. Summary

We have developed a linearized feedback control algorithm for a PDE model of a re-entrant production line, based on changing dispatch policies. Specifically, we choose a point, called the PPP, that separates high priority production steps...
downstream from low priority production steps upstream. Moving this point hence changes the speed at which product moves through the factory. We can therefore use the position of the PPP as an actuator to control the outflux of the production line.

The linearized feedback is based on the assumption that the system at the time of the control action is in steady state and the resulting steady state profile then moves with a constant velocity, making the PDE a linear transport equation.

To study the feedback control algorithm we analyzed the production planning problem for random demand sequences with fixed mean demand rate and a start rate equal to the mean demand. We compare the output of the controlled system to the demand sequences for the feedback controlled system, a heuristic for controlling the PPP and two static dispatch policies.

We find that the feedback control algorithm always is as good or better than all the other control policies. In the best case scenario, the linearized feedback control improves the production planning outcomes on average by about 35%.

4.1. Practical considerations and limitations

We showed that the steady state assumption is not always valid. Specifically, short re-planning intervals will not allow the system enough time to relax to steady state and hence lead to poor performance of the linearized feedback control. Similarly, very large re-planning intervals are not impacted by a change in dispatch policies, since over a long period of time, all wip in the factory will leave the factory regardless of the dispatch policy. Hence in that case, the dispatch policy is not a valid actuator of the production rate.

On a practical level, we show (i) that an optimal re-planning interval leading to a maximal improvement of the system performance is achieved at about 1/2 of the mean cycle time of the factory. This matches intuition: Dispatch policies works by re-arranging priority on the wip currently in the factory. Hence a re-planning interval of 1/2 of the mean cycle time gives us an operating point on the wip that is about 1/2 of the wip in the factory thus allowing control to follow demand fluctuations that are symmetric above and below the mean demand rate. (ii) Since dispatch policies re-allocate free capacity to improve production speed, any control based on it will deteriorate as the utilization increases. For very high utilization, there is no free capacity to re-allocate and hence the feedback control will not improve the system performance much any more (clearly stochastic variations still allow some improvements but they will necessarily be small).

Overall, this is a conceptual study: semiconductor production lines that only produce one product, if they exist at all, are rare. Similarly, the aggregation of production to a continuum model prevents any detailed planning for individual lots and hence we are not solving the production planning problem for a particular day, or a particular set of products.

4.2. Future work

This paper considered a “linearized” re-entrant factory by unraveling the re-entrant dynamics on a linearly increasing variable denoting the production stages or the degree of completion of a product. The only viable control in this case is through control of the input and the dispatch rules. There is one more possible control parameter in such systems: If there are multiple paths through the factory that allow for choices at bifurcations of the path, then the assignment of parts to particular paths becomes a control parameter. The impact of global vs. local and autonomous control for such production networks has been studied in [14]. Linking all three control parameters together will be an interesting challenge for future work.

Another open problem is the question of the stability boundaries for the feedback control law. While our simulations indicate the algorithm is stable for random demand sequences, we have no proven this and we have not studied the question of the stability of the feedback control under parameter changes and/or stochastic parameters. Since the feedback control is for a nonlinear, nonlocal PDE, theoretical proofs or stability are hard and remain an open question.

Finally, we studied the feedback law under demand variation sampling from a uniform distribution. The uniform distribution was chosen as a baseline, other distributions more suited to demand models could be chosen to study particular demand scenarios and their impact on the feedback control efficiency.

Acknowledgments

This work has been supported by DFG Cluster of Excellence Integrative production technology for high-wage countries, BMBF Kinetic models and optimization for production processes KinOpt, DAAD 55866082. D.A wishes to acknowledge travel support from KI-Net, NSF RNMS Grant No. 1107291.

References


