Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Modeling, Validation and Control of Manufacturing Systems

E. Lefeber, R.A. van den Berg and J.E. Rooda

***Abstract*— In this paper we elaborate on the problem of supply chain control in semiconductor manufacturing. First, we introduce the problem. Next, we propose the use of Effective Processing Times (which can be measured from real factory data) to arrive at 'simple' discrete-event models for manufacturing systems. We explain why existing models can not be used for solving the problem and explain the need for PDE-models that consider the flow of products as a compressible fluid flow. Next, we present a validation study in which we compare the response of the currently available PDE-models to the results of discrete-event simulation. We conclude the paper by analytically deriving a controller that solves the ramp-up problem. The resulting controller is often used in practice.**

I. Introduction

The dynamics of manufacturing systems has been a subject of study for several decades [1], [2]. Even though understanding the dynamics of manufacturing systems is a challenging problem, studying the overall dynamics of a network of interacting manufacturing systems is even more challenging. A network of suppliers that produce goods, both for one another and for generic customers, is also called a supply chain. A simple example of a semiconductor manufacturing supply chain is given in Fig. 1.

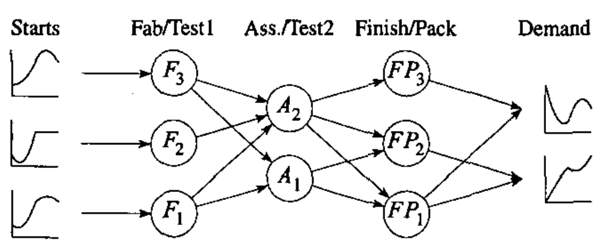

Fig. 1. A simple semiconductor manufacturing supply chain.

In this figure, F_1, F_2, and F_3 denote wafer fabs, in which wafers are being produced, containing hundreds to thousands of integrated circuits (ICs) on its surface. Due to, among others, the large number of process steps, the re-entrant nature of the process flow, and the advanced process technologies, the fabrication of wafers is a complex manufacturing process. A typical flow time for a wafer fab is in the order of two months. That is, once a bare silicon wafer enters the manufacturing system, it typically takes about two months for the wafer to be completed.

E. Lefeber, R.A. van den Berg and J.E. Rooda are with the Systems Engineering Group, Department of Mechanical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands [A.A.J.Lefeber,R.A.v.d.Berg,J.E.Rooda]@tue.nl

Finished wafers are moved to an Assembly/Test facility, where individual chips are cut out of the wafer and each separated IC is assembled. Typical flow times for the manufacturing systems A_1 and A_2 is in the order of ten days. Finally, the chips are packaged in FP_1, FP_2, FP_3, and can be shipped to customers. This takes in the order of five days.

The control of this supply chain is one of the problems the semiconductor industry currently faces. The fact that flow times are large and nonlinearly dependent on the load is one of the most difficult aspects in this problem. Notice that, even though the flow time of a wafer fab is in the order of two months, the pure processing time of a wafer is less than two weeks. That is, if a wafer enters an empty wafer fab, it takes less than two weeks for the wafer to be completed.

We are interested in solving the following problem: given a certain time-varying demand and the current state of the system, when to start how many new products (for each wafer fab) and how to coordinate the network flows. For addressing this question, we first need valid computationally feasible models that describe the dynamics of a manufacturing system and incorporate both throughput and flow time.

In this paper, we propose in Section II to first build a discrete-event model which describes the manufacturing system under consideration, using Effective Processing Times that can be estimated from real factory data [3]. In Section III we quickly review the models that have been used in the literature on modeling and control of manufacturing systems, we discuss why a new class models, PDE-models, is needed, and we give an overview of the PDE-models currently available. In Section IV we present a validation study of currently available PDE-models, in Section V we consider the control of one of these PDE-models, and Section VI concludes the paper.

II. Effective Processing Times

Several factors contribute to the flow time of lots in a manufacturing system, ranging from processing time, transport time, and variable availability of resources, to non-product lots, batching, setups, lots on hold, and rework lots. In semiconductor manufacturing industry it is common practice to build detailed discrete-event models incorporating all of the mentioned effects. One of the major disadvantages of these large models is their computational complexity: evaluating each what-if scenario can take hours for a typical semiconductor facility. A considerable reduction in model complexity can be achieved by considering Effective Processing Times (EPTs) as a conceptual way

0-7803-8335-4/04/$17.00 ©2004 AACC

of thinking to describe the combined influence of multiple sources of variability [3].

According to [2], the Effective Processing Time is the time seen by a lot from a logistical point of view. In other words, it is the time a lot actually was or could have been in process at a workstation. This idea can most clearly be illustrated by means of a simple example.

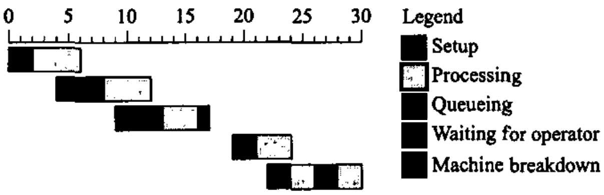

Fig. 2. Gantt chart of a workstation consisting of a buffer and one machine.

In Fig. 2, the Gantt chart of a workstation is depicted. The first lot arrives at $t = 0$ at this workstation. After the equipment has been setup, at $t = 2$ processing of the lot starts and is completed at $t = 6$. The second lot arrives at $t = 4$, but has to wait (since the first lot is being processed). Even though the equipment becomes available at $t = 6$, when the first lot leaves the workstation, it takes until $t = 7$ before the second lot starts processing, as only then an operator is available. At $t = 12$ the second lot is completed. The third lot, which arrives at $t = 9$, after a setup finishes processing at $t = 16$. However, due to the fact that no operator is available, the lot leaves the equipment only at $t = 17$. For the fifth lot processing starts at $t = 24$, but at $t = 26$ a machine breakdown occurs. The machine is up again at $t = 28$ and finishes the fifth lot at $t = 30$.

How can we determine the Effective Processing Times of this equipment? In order to do so we take the perspective of a lot. The first lot arrives at $t = 0$ at an empty workstation. According to the lot, processing therefore starts at $t = 0$ and finishes at $t = 7$. Since the lot is not aware of what is involved in the processing of it, this is what the lot effectively experiences as processing time. The second lot arrives at $t = 4$ at a busy workstation. Therefore, it has to wait (which is a common experience for lots). However, at $t = 6$ the equipment becomes available, so according to the second lot its effective processing starts at $t = 6$. From the perspective of the second lot it can very well be part of its processing that an operator first does something else before putting the lot on the equipment. Using similar reasoning we can arrive at the EPTs of all five lots, as depicted in Fig. 3.

Notice that the Effective Processing Times are given by the time that a lot was, or could have been, in process. Furthermore, the only information needed for determining the EPTs of an equipment consists of the arrival and departure times of the lots processes by the equipment. As this data is commonly collected at semiconductor manufacturing systems, EPT-distributions of workstations can be determined from real factory data [3].

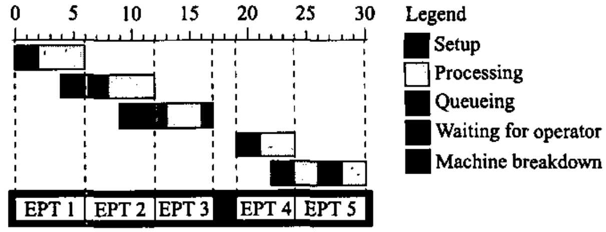

Fig. 3. EPT realizations at workstation.

These EPT-distributions capture not only the theoretical processing time, but also setup time, breakdown, operator availability, and all other operational times due to variability effects. This implies that EPT-distributions can be used for obtaining a so-called EPT-meta-model. That is, a discrete-event queueing model, consisting only of the equipments with their processing times drawn from their corresponding EPT-distributions. Since all variability has been incorporated in the EPT-distributions, there is no need for including all kinds of variability effects in the discrete-event model, like failure-behavior of machines, personnel behavior, etc. As a result, considerable reduction in the complexity of discrete-event simulation models for a semiconductor wafer fab can be achieved, while still yielding reasonable estimates for throughput and flow time.

The algorithms for determining EPT-realizations from real factory data as described in [3] can only be applied to equipment that processes single lots. Algorithms for equipment that processes lots in batches can be found in [4].

III. Model

Even though using an EPT-meta-model can considerably reduce the complexity of discrete-event models for manufacturing systems, the overall model of a semiconductor wafer fab is still unsuitable for dealing with the problem addressed in the introduction. A wafer fab consists of more than 100 machines and each wafer needs to undergo more than 100 processing steps. In the simple supply chain as presented in Fig. 1 three of such models need to run in parallel, together with several other models describing the dynamics of the other factories. Clearly, using discrete-event simulation for studying the supply chain control problem is computationally unfeasible. This also holds for the discrete-event models as studied by Ramadge and Wonham [5], since all possible states need to be considered in which a manufacturing system can be.

A second class of models available in the literature are models based on relations from queueing theory, see e.g., [6], [7]. Although these results give valuable insight into steady-state behavior of manufacturing lines, a disadvantage is that only steady state is concerned. No dynamic relations are available. Therefore, these models cannot be used for studying the supply chain control problem mentioned in the introduction.

A third class of models available in the literature are the so-called fluid models, in which the number of products is considered to be a continuous variable. Examples of these models are the flow model as initiated by Kimemia and Gershwin [8] for modeling failure-prone manufacturing systems, the fluid models or fluid queues as proposed by queueing theorists [9], or the stochastic fluid model as introduced by Cassandras et al. [10]. In these models, each buffer is modeled using the observation that the rate of change of the buffer contents is given by the difference between the rate at which lots enter and leave the buffer. Unfortunately, these models are only throughput oriented. The nonlinear relation between throughput and flow time is not incorporated in these models. As a result, a property of these models is that any feasible throughput can be achieved by means of zero inventory. Also, in case one feeds lots to an initially empty factory, according to these models lots will immediately come out of the manufacturing system, which in practice does not happen. Since large flow times play a crucial role in the supply chain control problem for the semiconductor industry, these fluid models can not be used either.

Recently, a new class of models for manufacturing systems has been introduced [11], [12], [13]. In these models, the flow of products through a manufacturing system is modeled in a similar way as the flow of cars on a highway. Not only is the number of lots assumed to be continuous, also the position of a lot in the manufacturing system is assumed to vary continuously.

Let $t \in \mathbb{R}^+$ denote the time and let $x \in [0,1]$ denote the position in the manufacturing line (the degree of completion). The behavior of lots flowing through the manufacturing line can now be described by three variables that vary with time and position: flow $u(x,t)$ measured in unit lots per unit time, density $\rho(x,t)$ measured in unit lots per unit completion and speed $v(x,t)$ measured in unit completion per unit time. First, we observe that flow is the product of density and speed:

$$u(x,t) = \rho(x,t)v(x,t). \tag{1}$$

Second, assuming no scrap, the number of products between any two locations x_1 and x_2 ($x_1 < x_2$) needs to be conserved at any time t, i.e., the change in the number of products between x_1 and x_2 is equal to the flow entering at x_1 minus the flow leaving at x_2:

$$\frac{\partial}{\partial t}\int_{x_1}^{x_2} \rho(x,t)\,\mathrm{d}x = u(x_1,t) - u(x_2,t),$$

or in differential form:

$$\frac{\partial \rho}{\partial t}(x,t) + \frac{\partial u}{\partial x}(x,t) = 0. \tag{2}$$

The two relations (1) and (2) are basic relations that any model must satisfy. As we have three variables of interest, a third relation is needed. For this third relation, several choices can be made. So far, the following models have been proposed:

Model 1 (Single queue I, cf. [11]): Relations (1), (2) together with

$$v(x,t) = \frac{\mu}{1+\int_0^1 \rho(s,t)\,\mathrm{d}s}, \tag{3}$$

where $\mu > 0$ is a constant representing the processing rate of the workstation.

Model 2 (Single queue II, cf. [11]): Relations (1),(2) together with

$$\frac{\partial \rho v}{\partial t}(x,t) + \frac{\partial \rho v^2}{\partial x}(x,t) = 0, \tag{4}$$

and the additional boundary condition

$$\rho v^2(0,t) = \frac{\mu \cdot \rho v(0,t)}{1+\int_0^1 \rho(s,t)\,\mathrm{d}s}, \tag{5}$$

where $\mu > 0$ again denotes the processing rate of the workstation.

Model 3 (Re-entrant I, cf. [12]): Relations (1), (2) together with

$$v(x,t) = v_0\left(1 - \frac{1}{L_{\max}}\int_0^1 \rho(s,t)\,\mathrm{d}s\right), \tag{6}$$

where $v_0 > 0$ is a constant representing the maximal speed that can be achieved (i.e., $1/v_0$ denotes the theoretical minimal flow time), and $L_{\max} > 0$ is a constant representing the maximal number of lots that can be in the manufacturing system.

Model 4 (Re-entrant II, cf. [12]): Relations (1), (2) together with (4), and the additional boundary condition

$$\rho v^2(0,t) = v_0\left(1 - \frac{1}{L_{\max}}\int_0^1 \rho(s,t)\,\mathrm{d}s\right)\rho v(0,t), \tag{7}$$

where v_0 and $L_{\max}$ are the same as in (6).

Model 5 (m identical machines, cf. [13]): Relations (1), (2) together with

$$v(x,t) = \frac{\mu}{m+\rho(x,t)},$$

where $m > 0$ denotes the number of machines, and $\mu > 0$ denotes the processing rate of each workstation.

All five models have as a boundary condition

$$\rho v(0,t) = \lambda(t),$$

where $\lambda(t)$ denotes the inflow to the manufacturing system (the lot start rate) in unit lots per unit time.

When we compare the PDE-models 1–5 to the other models available in the literature, the PDE-models are the only ones that are *computationally feasible*, describe the *dynamics* of a manufacturing system and incorporate *both throughput and flow time*. Furthermore, as we illustrate in Section V, we can study the boundary control problem for PDE-models, yielding an answer to the question how to feed lots to a manufacturing system. The only question remaining is: are the models 1–5 valid models?

IV. Validation study

In the previous section we discussed that discrete-event models of manufacturing systems, as well as models from queueing theory, are not suitable for addressing the supply chain control problem as mentioned in the introduction. Nevertheless, these are well-accepted models in the analysis of manufacturing systems. Therefore, queueing theory and discrete-event simulation can be be used for validating the models 1–5.

When we consider the supply chain in Fig. 1, two typical manufacturing systems can be considered. On the one hand we have the factories F_1, F_2, and F_3, which have a re-entrant nature, on the other hand we have the factories A_1, A_2, FP_1, FP_2, and FP_3, which have the nature of a line of workstations. Therefore, we define two manufacturing systems:

System 1: A line consisting of 15 identical workstations. Lots visit the workstations according to the following recipe: 1-2-3-4-5-6-7-8-9-10-11-12-13-14-15. This is an 'ordinary' manufacturing line.

System 2: Consider a line consisting of five identical workstations. Lots visit the workstations according to the following recipe: 1-2-3-4-5-1-2-3-4-5-1-2-3-4-5. Since each lot re-enters the system twice, this is a re-entrant manufacturing line.

We assume that each workstation consists of an infinite buffer, which operates under a FIFO policy (First In First Out), and a single machine whose Effective Processing Times are drawn from an exponential distribution with mean 1. If we furthermore assume that lots arrive to the manufacturing system according to a Poisson process with an arrival rate λ, we can easily derive the following steady state properties by means of queueing theory:

- For System 1, the mean number of lots equals $\frac{\lambda}{1-\lambda}$ in each workstation, resulting in a mean number of $\frac{15\lambda}{1-\lambda}$ lots in the system. Furthermore, the mean flow time of lots for System 1 is $\frac{15}{1-\lambda}$. Translated into PDE-terms we have
$$\rho(x,t) = \frac{15\lambda}{1-\lambda}, \qquad v(x,t) = \frac{1-\lambda}{15}. \tag{8}$$
- For System 2, the mean number of lots equals $\frac{3\lambda}{1-3\lambda}$ in each workstation, resulting in a mean number of $\frac{15\lambda}{1-3\lambda}$ lots in the system. Furthermore, the mean flow time of lots for System 2 is $\frac{15}{1-3\lambda}$. Translated into PDE-terms we have
$$\rho(x,t) = \frac{15\lambda}{1-3\lambda}, \qquad v(x,t) = \frac{1-3\lambda}{15}. \tag{9}$$

From (8) we obtain, by eliminating λ, that in steady state
$$v(x,t) = \frac{1}{15+\rho(x,t)} = \frac{1}{15+\int_0^1 \rho(s,t)\,\mathrm{d}s}. \tag{10}$$
Similarly, from (9) we obtain that in steady state
$$v(x,t) = \frac{1}{15+3\rho(x,t)} = \frac{1}{15}\left(1 - \frac{\int_0^1 \rho(s,t)\,\mathrm{d}s}{\frac{5}{1-3\lambda}}\right). \tag{11}$$

When we compare (11) with (6) and (7), we notice that in order for models 3 and 4 to be valid in steady state, we need $L_{\max} = \frac{5}{1-3\lambda}$, where λ denotes the steady state arrival rate. Therefore, the re-entrant models 3 and 4 are not likely to be 'globally' valid for re-entrant manufacturing systems. In the best case they are valid 'locally' around a certain steady state. On the other hand, any manufacturing system can contain only a finite number lots, arguing the validity of a queueing model with infinite buffers.

From (10) we obtain that the models 1, 2, and 5 are valid in steady state, provided that in (3) and (5) we replace the denominator $1+\int_0^1 \rho(s,t)\,\mathrm{d}s$ with $15+\int_0^1 \rho(s,t)\,\mathrm{d}s$, which is consistent with the results in [11]. In [11] a single queue is assumed. If, instead, we assume a line of 15 workstations the mentioned modification of (3) and (5) results.

Next, we can use discrete-event models of System 1 and System 2 to study the dynamics of the proposed PDE-models. Starting with an initially empty system, we performed experiments where lots arrive according to a Poisson process with a mean arrival rate λ. During an experiment we collected at the times $t = 1,2,3,\dots$ the number of lots in each workstation as well as the number of lots that has been completed by the system. In order to guarantee a 99% confidence interval with a relative width of less than 0.01 for *each* measurement, experiments have been repeated 1.000.000 times. We averaged all data, resulting in the average number of lots in each workstation, as well as the number of lots that has been completed by the system, at each time-instant. This we did for both System 1 and System 2, where we chose the arrival rate such that the steady-state utilization of the workstations was respectively 25%, 50%, 75%, 90%, and 95% (so $\lambda = 0.25, 0.5, 0.75, 0.90, 0.95$ for System 1 and $\lambda = 0.08333, 0.16667, 0.25, 0.3, 0.31667$ for System 2). Clearly, these experiments provide more data than can be presented in this paper. The interested reader is referred to [14] for more results. Here we present some general findings.

The first results we present are for System 1 with an arrival rate of $\lambda = 0.25$. Fig. 4 presents the evolution of the total number of lots in the system as a function of time. The black solid line describes the (averaged) result of the

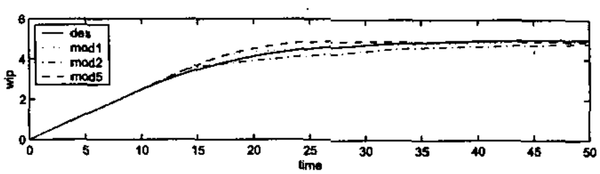

Fig. 4. Number of lots in System 1 for utilization of 25%.

discrete-event simulations. The magenta dotted line, the red dash-dotted line, and the blue dashed line describe the result according to respectively Model 1, Model 2, and Model 5. In Fig. 4 we see that initially the total number of lots in the line linearly increases. This is due to the fact that lots are only entering the system and it takes a while before lots start coming out. Also, we see that all models predict that in steady state five lots are in the system. This is as expected. When we closely look at Fig. 4 we see that around $t = 10$ the graph of the discrete-event simulation bends off from the PDE-graphs, from which we can conclude that the moment at which the first lot leaves the factory is overestimated by the PDE-models. That is, according to the discrete-event simulation this should happen earlier. Also, we see that after $t = 40$ all three PDE-models underestimate the number of products in the system. Therefore, all PDE-models predict that the system is later in steady state than according to the discrete-event simulation.

The differences in behavior become more clear when we consider the development of the density over time. This can be made most clear by means of a movie, for which the reader is referred to [14]. In Fig. 5 the most important parts of the behavior are captured. The figure presents respectively $\rho(0,t)$, $\rho(0.5,t)$ and $\rho(1,t)$, again for the discrete-event model, Model 1, Model 2, and Model 5. For the discrete-event system we assume the density to be piecewise constant at intervals of width $\frac{1}{15}$. When looking

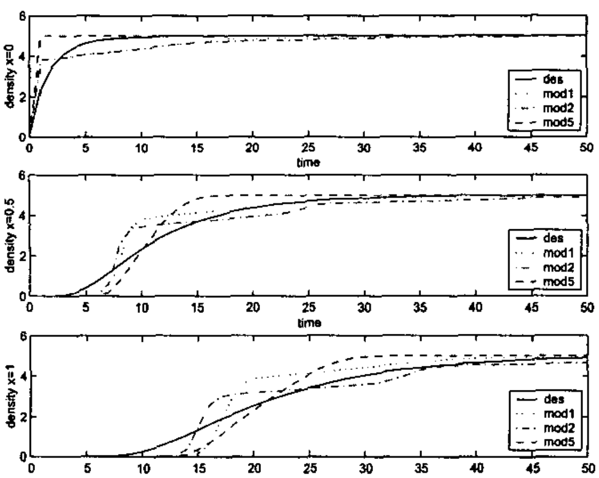

Fig. 5. Densities at $x = 0$, $x = 0.5$ and $x = 1$ for utilization of 25%.

at the first graph, we see that the behavior of Model 1 and Model 2 almost coincide. All three models predict a quicker raise of the density than the discrete-event model predicts. If we look at the graph of $\rho(0.5,t)$ we see that initially the PDE-models underestimate the growth of the density, around $t = 7$ the PDE-models show a strong increase in the density, resulting in an over-estimation of the density. Similar behavior can be observed for $\rho(1,t)$.

The second results we present are for System 2 with an arrival rate of $\lambda = 0.08333$. Fig. 6 presents the evolution of the total number of lots in the system as a function of time. In addition to the lines from the previous two figures, the green and cyan solid line represent the output of Model 3 and Model 4 respectively. The third equations in models 1, 2 and 5 have been modified according to the difference

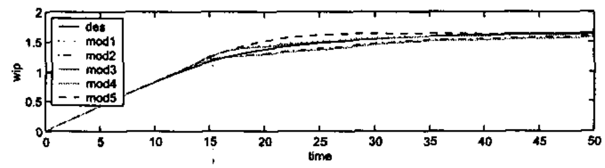

Fig. 6. Number of lots in System 2 for utilization of 25%.

between (10) and (11).

For the re-entrant case we can make similar remarks as for System 1. Furthermore, a close resemblance between Model 1 and Model 3 can be noticed, as well as a close resemblance between Model 2 and Model 4.

To conclude this validation study, we remark that Model 1 has a uniform velocity through the whole factory. As a result, lots at the end of the line are influenced by lots in the beginning of the line. For System 1 this is an undesirable property. If initially the manufacturing system is non-empty, increasing the influx results in an initially decreasing outflux. In an actual manufacturing line this does not happen. Clearly, more accurate models are needed. Recently, a new model has been proposed in [15]. It would be interesting to include this model in the validation study.

V. CONTROL

Even though the validation study of the previous section showed that current PDE-models do not describe the dynamics of manufacturing systems in the same way as discrete-event simulations do, they do provide several advantages. First of all, simulating a PDE-model takes in the order of seconds, whereas simulating a discrete-event model takes in the order of hours. However, even more important is that fact PDE-models can used for analytically deriving control strategies.

In case we have a PDE-model for each manufacturing system of the supply chain in Fig. 1, the problem of determining when to start how many products (for each wafer fab) and how to coordinate the network flows, simply becomes a (left-)boundary control problem for PDEs. To illustrate how such a problem might be tackled, we consider the problem of ramping up a line, like System 1.

From the previous section we know that models 1–4 have undesirable properties. Therefore, consider Model 5, even though it clearly also has its shortcomings:

$$\frac{\partial \rho(x,t)}{\partial t} + \frac{\mu m}{(m+\rho(x,t))^2}\frac{\partial \rho(x,t)}{\partial x} = 0, \qquad (12)$$

with boundary condition $\rho(0,t) = \rho_0(t)$, as input.

Assume that we want this system to converge to the steady state

$$\rho(x,t) = \rho_{ss} = \frac{m\lambda_{ss}}{1-\lambda_{ss}}.$$

Consider the Lyapunov function candidate

$$V = \frac{2}{3\mu m}\int_0^1 \left[(m+\rho(s,t))^3 - (m+\rho_{ss})^3\right]^2 ds, \qquad (13)$$

which is positive for all $\rho(x,t) \neq \rho_{ss}$. Along the dynamics (12) we have

$$\begin{aligned}
&\frac{\partial}{\partial t}\left[(m+\rho)^3-(m+\rho_{ss})^3\right]^2\\
&=2\left[(m+\rho)^3-(m+\rho_{ss})^3\right]\cdot 3(m+\rho)^2\cdot\frac{\partial \rho}{\partial t}\\
&=-6\cdot\left[(m+\rho)^3-(m+\rho_{ss})^3\right]\cdot(m+\rho)^2\cdot\frac{\mu m}{(m+\rho)^2}\cdot\frac{\partial \rho}{\partial x}\\
&=-6\mu m\cdot\left[(m+\rho)^3-(m+\rho_{ss})^3\right]\cdot\frac{\partial \rho}{\partial x}.
\end{aligned}$$

Therefore, differentiating the Lyapunov function candidate (13) along the dynamics (12) results in

$$\begin{aligned}
\dot{V}&=-\int_0^1 4\left[(m+\rho(x,t))^3-(m+\rho_{ss})^3\right]\frac{\partial \rho}{\partial x}\,\mathrm{d}x\\
&=[m+\rho_0(t)]^4-[m+\rho_1(t)]^4+4[m+\rho_{ss}]^3[\rho_1(t)-\rho_0(t)],
\end{aligned}$$

where $\rho_1(t)=\rho(1,t)$.

If we want to reach the desired steady state as quickly as possible, we should try to minimize $\dot{V}$ by a proper choice of $\rho_0(t)$. It is easy to verify that we minimize $\dot{V}$ by taking

$$\rho_0(t)=\rho_{ss}. \tag{14}$$

As a result, we obtain

$$\begin{aligned}
\dot{V}&=[m+\rho_{ss}]^4-[m+\rho_1(t)]^4+4[m+\rho_{ss}]^3[\rho_1(t)-\rho_{ss}]\\
&=-\frac{1}{3}[\rho_1(t)-\rho_{ss}]^4-\frac{2}{3}[\rho_1(t)+2\rho_{ss}+3m]^2[\rho_1(t)-\rho_{ss}]^2,
\end{aligned}$$

which is negative for $\rho_1(t)\neq\rho_{ss}$.

We establish the following result:

Proposition 1: Consider the system (12) together with the input (14). Then we have

$$\lim_{t\to\infty}\rho(x,t)=\rho_{ss}.$$

Furthermore, the choice (14) is the input that achieves the goal the quickest.

Notice that the boundary control (14) does not only achieve stabilization to the desired steady state the quickest, it is also a very simple input to be applied. Actually, this input is common practice when ramping up semiconductor manufacturing systems.

VI. Conclusions

In this paper we discussed the problem of controlling a semiconductor manufacturing supply chain, i.e., given a certain time-varying demand and the current state of the system: when to start how many products for each factory and how to coordinate the network flows.

Several factors contribute to the flow time of lots in a manufacturing system, ranging from processing time, transport time, and variable availability of resources, to non-product lots, batching, setups, lots on hold, and rework lots. Instead of including all these factors into a complex large discrete-event model, we proposed to use Effective Processing Times for capturing all variability, yielding a much simpler discrete-event queueing model.

Unfortunately, discrete-event queueing models of semiconductor manufacturing systems are still too large to be able to successfully address the supply chain control problem. Also other established models such as queueing theory and fluid queues of flow models are unsuitable. Therefore, we introduced PDE-models in which the flow of products is considered as a compressible fluid flow. This new class of models is *computationally feasible*, describes the *dynamics* of a manufacturing system, and incorporates *both throughput and flow time*, and can be used for addressing the supply chain control problem.

Next, the currently available PDE-models have been reviewed and validated by means of queueing theory and discrete-event simulation. A need for more accurate models was made clear.

Finally, a ramp-up control problem has been studied using one of the available PDE models. It turned out that, in order to reach full production in the shortest time, one can best feed the manufacturing system at the desired steady-state rate. This is a simple control action as often used in practice.

References

[1] J.W. Forrester, *Industrial Dynamics*. Cambridge, MA, USA: MIT Press, 1961.

[2] W.J. Hopp and M.L. Spearman, *Factory Physics*, 2nd ed. New York, USA: Irwin/McGraw-Hill, 2000.

[3] J.H. Jacobs, L.F.P. Etman, E.J.J. v. Campen, and J.E. Rooda, "Characterization of the operational time variability using effective processing times," *IEEE Transactions on Semiconductor Manufacturing*, vol. 16, no. 3, pp. 511–520, August 2003.

[4] P.P. v. Bakel, J.H. Jacobs, L.F.P. Etman, and J.E. Rooda, "Quantifying variability of batching equipment using effective process times." 2004, submitted to IEEE Transactions on Semiconductor Manufacturing.

[5] P.J. Ramadge and W.M. Wonham, "Supervisory control of a class of discrete-event systems," *SIAM Journal on Control and Optimization*, vol. 25, pp. 206–230, 1987.

[6] J.A. Buzacott and J.G. Shantikumar, *Stochastic Models of Manufacturing Systems*. Englewood Cliffs, New Jersey, USA: Prentice Hall, 1993.

[7] R. Suri, *Quick Response Manufacturing: A Companywide Approach to Reducing Lead Times*. Portland, Oregon, USA: Productivity Press, 1998.

[8] J. Kimemia and S.B. Gershwin, "An algorithm for the computer control of a flexible manufacturing system," *IIE Transactions*, vol. 15, no. 4, pp. 353–362, December 1983.

[9] J.M. Harrison, *Brownian Motion and Stochastic Flow Systems*. New York: John Wiley, 1995.

[10] C.G. Cassandras, Y. Wardi, B. Melamed, G. Sun, and C. Panayiotou, "Perturbation analysis for on-line control and optimization of stochastic fluid models," *IEEE Transactions on Automatic Control*, vol. 47, no. 8, pp. 1234–1248, August 2002.

[11] D. Armbruster, D. Marthaler, and C. Ringhofer, "Kinetic and fluid model hierarchies for supply chains," *SIAM Journal on Multiscale Modeling and Simulation*, vol. 2, no. 1, pp. 43–61, 2004.

[12] ——, "Modeling a re-entrant factory," 2004, submitted to Operations Research. [Online]. Available: http://math.la.asu.edu/~chris.

[13] E. Lefeber, "Nonlinear models for control of manufacturing systems," in *Nonlinear Dynamics of Production Systems*, G. Radons and R. Neugebauer, Eds. Berlin, Germany: John Wiley, 2004, ch. 5, pp. 69–81.

[14] ——, "Movies containing results of the validation studies," 2004. [Online]. Available: http://se.wtb.tue.nl/~lefeber/acc04/movies/.

[15] D. Armbruster and C. Ringhofer, "Thermalized kinetic and fluid models for re-entrant supply chains," 2004, submitted to SIAM Journal on Multiscale Modeling and Simulation. [Online]. Available: http://math.la.asu.edu/~chris.