

KINETIC AND FLUID MODEL HIERARCHIES FOR SUPPLY CHAINS*

D. ARMBRUSTER[†], D. MARTHALER[‡], AND C. RINGHOFER[†]

Abstract. We present a model hierarchy for queuing networks and supply chains, analogous to the hierarchy leading from the many body problem to the equations of gas dynamics. Various possible mean field models for the interaction of individual parts in the chain are presented. For the case of linearly ordered queues the mean field models and fluid approximations are verified numerically.

Key words. supply chains, gas dynamics, kinetic equations

AMS subject classifications. 65N35, 65N05

DOI. 10.1137/S1540345902419616

1. Introduction. Prevalent models for complex manufacturing systems and supply chains are based either on discrete event simulations or on fluid models. They all describe the following generic situation: A sequence of products or parts arrives in the system at arbitrary arrival times a_n , $n = 1, 2, \dots, N$. They pass from one node (supplier) in the system to the next at rates determined by the orders from the downstream supplier and leave the system at exit times e_n , $n = 1, 2, \dots, N$. The object of modeling this system is to compute the dependence of the exit times on the arrival times and consequently the load of the system or “work in progress,” at any time t , given by $W(t) = \sum_{n=1}^N \chi_{[a_n, e_n]}(t)$, where $\chi_{[a, e]}$ denotes the usual indicator function.

Existing approaches to this generic problem range from the most “microscopic” description, namely discrete event simulators (see [5] for an introduction and overview), to the most macroscopic description, namely rate equations, or fluid models as they are called in the industrial engineering literature (see [1], [6]). Discrete event simulators model each individual part by setting it into a queue, or subqueue in a more complex system, and advancing its position in the queue at discrete time intervals given by the processing time of the components of the system. Rate equations, on the other hand, consider only the length W of the queue and model it by an ordinary differential equation of the form $W'(t) = F^{in} - F^{out}$, where F^{in} denotes the known influx and F^{out} is modeled more or less heuristically by rules describing the policy of the chain. More complicated systems can then be modeled by breaking them down into individual queues. For instance, a system consisting of a simple chain of J individual stations can be modeled according to

$$W'_j(t) = F_j^{in} - F_j^{out}, \quad F_j^{in} = F_{j-1}^{out}, \quad j = 2, \dots, J, \quad W(t) = \sum_{j=1}^J W_j(t).$$

The challenge then becomes to find models for the outflux functions F_j^{out} such that the individual queue lengths W_j do not become negative. This leads, in general, to

*Received by the editors December 12, 2002; accepted for publication (in revised form) September 4, 2003; published electronically November 25, 2003. This research was supported by NSF grant DMS-0204543.

<http://www.siam.org/journals/mms/2-1/41961.html>

[†]Department of Mathematics, Arizona State University, Tempe, AZ 85287-1804 (armbruster@asu.edu, ringhofer@asu.edu).

[‡]Department of Mathematics, Duke University, Durham, NC 27708-0320 (daniel@math.ucla.edu).

continuous versions of linear programming problems [19]. Obviously, rate equations are less exact than discrete event simulators but computationally more efficient since their complexity does not depend on the number N of parts to be processed. Therefore, they are able to handle more complex systems and can be used for more complex tasks such as optimization of the system or the study of its long time behavior in the framework of dynamical systems.

The present paper is concerned with a similar sort of model hierarchy, starting from a more or less exact microscopic description in the form of Newton equations for individual parts and leading to an approximate macroscopic model, namely hyperbolic conservation laws. This approach is analogous to the methodology used in modeling flows of vehicles on a highway, where fluid equations for the density of vehicles are derived from simple rules governing the interaction of individual cars [12], [13], [16], [17], [18]. The success of this kind of methodology depends essentially on the accuracy of the rules governing the behavior of individual parts (or cars) and of the validity of the continuum approximation. For traffic flow models these questions are still subject to debate [9], [4]. This paper tries to address the following questions:

- What microscopic model for the movement of parts in the supply chain allows for the formulation of an effective single part model for the many body problem; i.e., what are possible mean field theories?
- What is the appropriate framework for the continuum limit; i.e., what is the error introduced by replacing an ensemble of parts by a continuum density?

These questions are of interest since continuum models are much more computationally efficient when dealing with a large number of parts.

We will use an approach which is analogous to the one used in gas dynamics, i.e., starting from classical many body problems and culminating in the compressible Euler equations. The resulting macroscopic model, which will be a system of conservation laws, is computationally more complex than the simple rate equation but provides significantly more information. However, its computational complexity also does not depend on the number of parts. Therefore, it can be viewed as a compromise between rate equations and discrete event simulators.

The key underlying idea is to introduce an artificial continuous variable, called “degree of completion” x , which measures the position of a part in the system; i.e., a part enters the system at $x = 0$ and leaves the system upon arriving at $x = 1$. Denoting the position of part number n at time t with $x = z_n(t)$ and its velocity by v_n we compute the evolution of z_n by the initial value problem

$$(1.1) \quad z'_n(t) = v_n(t), \quad z_n(a_n) = 0, \quad n = 1, \dots, N.$$

The modeling part of this system now consists of defining the velocities v_n . In the most general case, we will assume a law of the form

$$(1.2) \quad v_n(t) = \vartheta_n(t, A, Z(\tau), V(\tau), \tau < t), \\ A = (a_1, \dots, a_N), \quad Z = (z_1, \dots, z_N), \quad V = (v_1, \dots, v_N);$$

i.e., in the most general case the dynamics of the system (1.1) will be given by the arrival times of all parts and all its history. The dependence of ϑ_n on Z and V is in general of a functional form, and therefore the computation of the velocities v_n might involve the evaluation of integrals along particle paths or solving differential equations. The corresponding Newton equations are derived by computing the time

evolution of the velocities and are of the form

(1.3)

$$(a) z'_n(t) = v_n, \quad (b) v'_n(t) = s_n(t, A, Z(\tau), V(\tau), \tau < t) = \frac{d}{dt} \vartheta_n(t, A, Z(\tau), V(\tau), \tau < t),$$

$$(c) z_n(a_n) = 0, \quad v_n(a_n) = \vartheta_n(a_n, A, Z(\tau), V(\tau), \tau < a_n), \quad n = 1, \dots, N.$$

The dependence of ϑ on the state variables is often given in a functional form, in the sense that actually the acceleration terms s_n are modeled and the velocity functions ϑ_n are given as solutions of the ordinary differential equations (1.3)(b).

In this paper we will derive a model hierarchy for the system described by (1.3) which is analogous to the one of classical many body physics. From the Newton equations (1.3) we will derive the corresponding kinetic equation for an effective single part phase space density function under a mean field assumption. We will then discuss several possible mean field approximations for simple $M/M/1$ queues (meaning Markovian arrivals and processing times) and more complex supply chains. Hyperbolic conservation laws are obtained from the moment equations for the kinetic problem in the usual manner, where we will use the closure assumption that parts overtaking each other is a rare event. This leads to the equivalent of the compressible Euler equations at zero temperature.

The computational advantage of this model lies, on one hand, in the fact that the computational complexity of simulating the system does not grow, and, as it turns out, actually decreases, with the number of parts considered. While the efforts to solve (1.3) and the computational work in a discrete event simulation grow with N the solution of the hyperbolic problem is independent of N . On the other hand, the fluxes do not have to be artificially constrained, as is the case for rate equations, in order to avoid negative queue lengths.

This paper is organized as follows. In section 2 we derive the general form of the kinetic initial boundary value problem for the effective single part density function. This problem will still need some form of mean field approximation to be consistently solvable. We will discuss such mean field approximations in section 3. Given the kinetic initial boundary value problem, we use a standard result in section 4 that, in the absence of caustics [14], the kinetic problem can be replaced by the zero temperature Euler equations. In order to arrive at a real continuum model we consider a smoothed version of the influx boundary condition and give an estimate of how the smoothed version approximates the original model.

2. The kinetic model. As a first step towards a more macroscopic description of the many body model (1.3) we derive an equivalent kinetic formulation; i.e., we define the effective one body kinetic density function for (1.3) by

$$(2.1) \quad f(x, u, t) = \sum_{n=1}^N \delta(x - z_n(t)) \delta(u - v_n(t)) H(t - a_n), \quad t \geq 0, \quad 0 \leq x \leq 1,$$

where z_n, v_n is the solution of (1.3) and H denotes the usual Heaviside function, which has to be included since the particle (or part) number n exists only for $t > a_n$. We expect the kinetic density function f in (2.1) to satisfy the usual effective single body Liouville equation [8] with an inflow boundary condition. We derive the following initial boundary value problem for the kinetic density function f .

THEOREM 1. *The density function f defined by (2.1) satisfies the initial boundary value problem*

$$(2.2) \quad (a) \quad \partial_t f + u \partial_x f + \partial_u s = 0, \quad f(x, u, 0) = 0, \quad u f(0, u, t) = f^b(u, t),$$

weakly in x, u, t , where the field $s(x, u, t)$ and the boundary function $f^b(u, t)$ are given by

$$(b) \quad s(x, u, t) = \sum_{n=1}^{\infty} s_n(t, A, Z(\tau), V(\tau), \tau < t) H(t - a_n) \delta(x - z_n(t)) \delta(u - v_n(t)),$$

$$(c) \quad f^b(u, t) = \sum_{n=1}^N \delta(u - v_n(a_n, A, Z(\tau), V(\tau), \tau < a_n)) \delta(t - a_n).$$

Proof of Theorem 1. In order to compute the time evolution of the kinetic density f in (2.1) in a weak sense, we compute the derivative of a sufficiently smooth test function ψ along a particle path and obtain

$$\frac{d}{dt} \psi(z_n(t), v_n(t), t) = (\partial_t \psi + v_n \partial_x \psi + s_n \partial_u \psi)(z_n(t), v_n(t), t).$$

Integrating the above equation from $t = a_n$ to $t = \infty$, assuming that the test function ψ has compact support, gives

$$-\psi(0, v_n(a_n), a_n) = \int_{a_n}^{\infty} (\partial_t \psi + v_n \partial_x \psi + s_n \partial_u \psi)(z_n(t), v_n(t), t) dt,$$

and summation over the index n gives that

$$\begin{aligned} & - \int_0^{\infty} dt \int du \left[\psi(0, u, t) \sum_{n=1}^N \delta(u - v_n(a_n)) \delta(t - a_n) \right] \\ &= \int_0^{\infty} dt \int_0^{\infty} dx \int du \left[\sum_{n=1}^{\infty} H(t - a_n) \delta(x - z_n) \delta(u - v_n) (\partial_t \psi + u \partial_x \psi + s_n \partial_u \psi)(x, u, t) \right] \end{aligned}$$

holds for all sufficiently smooth compactly supported test functions ψ , which is the weak form of the initial boundary value problem

$$\partial_t f + u \partial_x f + \partial_u s = 0, \quad f(x, u, 0) = 0, \quad u f(0, u, t) = \sum_{n=1}^N \delta(u - v_n(a_n)) \delta(t - a_n),$$

and the function $s(x, u, t)$ is given by

$$s(x, u, t) = \sum_{n=1}^{\infty} s_n(t) H(t - a_n) \delta(x - z_n(t)) \delta(u - v_n(t)). \quad \square$$

Remark. The initial boundary value problem (2.2) can, of course, not be solved without the advance knowledge of the solution (Z, V) of the particle model. This is a consequence of (2.2) already being the equation for an effective single particle density.

To arrive at a solvable problem for f the acceleration field s and the boundary function f^b would have to be replaced by some mean field approximation of the form

$$s = \hat{s}(x, u, t, f), \quad f^b(u, t) = \hat{f}^b(u, t, f),$$

where \hat{s}, \hat{f}^b depend on the density function f in a certain way. (Usually this is a functional dependence.)

For the rest of this paper we will assume that the velocity models ϑ_n in (1.3)(b) are of a mean field form, i.e., that ϑ_n depends on the other parts z_ν , $\nu \neq n$, only through the density f . We set

$$(2.3) \quad s_n(t, A, Z(\tau), V(\tau), \tau < t) = S(t, A, z_n(t), v_n(t), f),$$

$$\vartheta_n(a_n, A, Z(\tau), V(\tau), \tau < a_n) = \vartheta^0(a_n, A, f),$$

where the dependence on the density f will, in general, be of a functional form. In this case, the acceleration term s and the boundary datum f^b in (2.2) are given by

$$(2.4) \quad (a) \quad s = s(x, u, t, f) = S(t, A, x, u, f)f,$$

$$(b) \quad f^b(u, t) = \delta(u - \vartheta^0(t, A, f)) \sum_{n=1}^N \delta(t - a_n).$$

For the rest of this paper we will assume the velocity model ϑ_n of this mean field form and consider the kinetic equation (2.2) with the acceleration field and the boundary condition by (2.4). Note that, other than in the usual Liouville and Boltzmann equation [8], we still allow the velocity model and the acceleration field s to depend on the arrival times A of the parts in the system.

Remark. We formally have defined the density function f for all values of the velocity u . If the particle model (1.3) is such that the velocities v_n always remain non-negative, then the kinetic density $f(x, u, t)$ will always be supported inside $\{u \geq 0\}$.

3. Velocity models. In this section we will discuss several possible models for the velocity function ϑ in (2.3). The first two deal with the case of simple linear queues with constant processing rates. This implies the following principles:

- Since the values of the degree of completion x between $x = 0$ and $x = 1$ have no special meaning (i.e., we do not specify what it means for a part to be half processed) the velocities can be set to constants.
- The exit time e_n of a part from the system, and therefore the constant velocity, is determined by the state of the system at the arrival time a_n .
- The parts leave the system in the same order in which they have arrived; i.e., moving with constant velocities, they cannot pass each other.

The first two of these principles imply that the function ϑ in (2.3) is dependent on A , the collection of arrival times, only, and therefore the acceleration field s is set to zero. We will first discuss an exact model, which is not of mean field form, since the boundary density f^b in (2.4) will still depend on all arrival times. Then we will give an approximate model which is based on Little's law and the statistical properties of Markov chains and is of a mean field form. The third model consists of a phenomenological approach to more complex supply chains, which takes into account the experimental observation that reentrant supply chains have a certain critical workload.

3.1. An exact model for queues. If we set the part velocities to constants, these constants have, because of the above principles, to be determined at the time the part enters the queue. Therefore, we can choose ϑ_n in (1.2) as

$$(3.1) \quad \vartheta_n(t, Z(\tau), V(\tau), \tau < t) = \vartheta_n(a_n, z_1(a_n), \dots, z_{n-1}(a_n), v_1, \dots, v_{n-1})$$

in this case, and the accelerating field s in (2.2) vanishes identically. The boundary datum f^b is then of the form

$$f^b(u, t) = \sum_{n=1}^N \delta(u - \vartheta_n(a_n, z_1(a_n), \dots, z_{n-1}(a_n), v_1, \dots, v_{n-1})) \delta(t - a_n).$$

Before proceeding further, we will discuss an important condition, namely, that neighboring particle paths cannot cross. The assumptions underlying our model imply that parts leave the system in the same order in which they have entered, i.e., that $z_n(t) < z_{n-1}(t)$ holds at least as long as $z_{n-1}(t) < 1$ holds.

LEMMA L1. *The particle paths will not cross, i.e., $z_n(t) < z_{n-1}(t)$ will hold as long as $z_{n-1}(t) < 1$, if and only if the sequence v_n , defined recursively by*

$$v_n = \vartheta_n(a_n, z_1(a_n), \dots, z_{n-1}(a_n), v_1, \dots, v_{n-1}),$$

satisfies

$$(3.2) \quad \frac{1}{v_{n-1}} - \frac{1}{v_n} < a_n - a_{n-1} \quad \forall n.$$

Proof of Lemma L1. At $t = a_n > a_{n-1}$, $z_n(a_n) = 0$ and $z_{n-1}(a_n) = v_{n-1}(a_n - a_{n-1}) > 0$ holds. Since z_n and z_{n-1} are linear functions of t , it is necessary and sufficient that $z_n(t) < 1$ holds for $t = e_{n-1}$, the exit time when $z_{n-1} = 1$ holds. From the linear form of z_{n-1} we compute the exit time e_{n-1} as $e_{n-1} = a_{n-1} + \frac{1}{v_{n-1}}$, giving the condition $v_n(a_{n-1} + \frac{1}{v_{n-1}} - a_n) < 1$. \square

We define e_n the exit time of part number n from the system. Denoting by T the time it takes to process one part, we obtain a simple recursion for the exit times in terms of the arrival times:

$$(3.3) \quad e_n = \begin{cases} e_{n-1} + T & \text{for } a_n < e_{n-1} \\ a_n + T & \text{for } e_{n-1} < a_n \end{cases} = T + \max\{e_{n-1}, a_n\}.$$

Formula (3.3) means that if the queue is not empty, i.e., $a_n < e_{n-1}$, part number n has to wait for number $n-1$, and then it takes an additional time T for it to be processed. On the other hand, if the queue is empty when part n arrives, i.e., $e_{n-1} < a_n$, it just takes a time T for it to be processed. If we set the velocity $v_n = \frac{1}{e_n - a_n}$ we obtain the velocity model

$$\vartheta_n^{ex}(z_{n-1}(a_n), v_{n-1}) = v_n = \frac{1}{T + \max\{0, \frac{1}{v_{n-1}} - a_n + a_{n-1}\}}.$$

The dependence on the previous arrival time a_{n-1} can be eliminated by using the fact that the part moves linearly in time; i.e., $z_{n-1}(a_n) = v_{n-1}(a_n - a_{n-1})$ holds. This gives

$$(3.4) \quad \vartheta_n^{ex}(a_n, z_{n-1}(a_n), v_{n-1}(a_n)) = \frac{1}{T + \max\{0, \frac{1 - z_{n-1}(a_n)}{v_{n-1}}\}}.$$

Equation (3.4) represents an exact velocity model for the queue, under the assumption of constant processing times, which is not in mean field form but obviously satisfies the no-path-crossing condition (3.2).

3.2. A mean field model for queues. We now turn to possible averaged models for the velocity function ϑ_n of a linear queue, which would allow us to formulate the kinetic problem (2.2) with acceleration and boundary terms of the form (2.4). In the classical many body problem, mean field approximations are derived from the assumption that for $N \rightarrow \infty$ individual particles become statistically independent, which, in an approximate scaling, yields a limiting expression for the accelerating field s in (2.2)(b) [8]. For the case of a simple queue, the field s vanishes anyway, and this approximation is needed for the boundary function f^b in (2.2)(c). The corresponding approximation for large values of N in the present model is given by Little's law and the stochastic properties of queuing systems [7], [11]. Little's law states that the average length of a queue is proportional to the average start rate and cycle time; i.e., $W = \tau\lambda$ holds, where W in our framework is given by $\int_0^1 \rho dx$, τ is the mean cycle time, and λ is the start rate. In an $M/M/1$ queue, meaning Markovian arrivals and processing rates, the mean cycle time τ is given by $\tau = \frac{1}{\mu - \lambda}$ with μ the processing rate of the queue. In our notation we have $\mu = \frac{1}{T}$. Thus we have the relations

$$(3.5) \quad (a) \quad W = \lambda\tau, \quad (b) \quad \tau = \frac{T}{1 - \lambda T}, \quad W(t) = \int_0^1 \rho(x, t) dx,$$

from which we eliminate λ and obtain $\tau = T(1 + W)$ for the average cycle time. The mean field approximation now consists of replacing the cycle time $\frac{1}{v_n}$ of the individual part number n by the mean cycle time τ which is computed from (3.5). This represents a form of quasi-steady state approximation which assumes that λ and W vary slowly during the time the part spends in the queue. This gives the model

$$(3.6) \quad v_n = \vartheta^{mf}(a_n, \rho) = \frac{1}{T(1 + W(a_n - 0))}, \quad W(t) = \int_0^1 \rho(x, t) dx.$$

(The superscript mf stands for mean field.) Note that the acceleration field s still vanishes, since ρ is evaluated only at the arrival time a_n . Also, we denote by $W(a_n - 0)$ the one-sided limit of W from the left, since W will take on only integer values. The kinetic initial boundary value problem (2.2) is then of the form

$$(3.7) \quad \partial_t f + u \partial_x f = 0, \quad f(x, u, 0) = 0, \quad u f(0, u, t) = \delta(u - \vartheta^{mf}(t, \rho)) \sum_{n=1}^N \delta(t - a_n).$$

To compare the mean field model (3.7) to the exact sequence (3.3) we have carried out the following numerical experiment. We choose 1100 random arrival times, where the first 200 parts arrive at a rate of 0.666 parts per processing time T and the next 800 parts arrive at a rate of 2 parts per processing time T . Nothing arrives for the next 200 time units, and then another 200 parts arrive at a rate of 1 part per processing time. Figure 1 shows the arrival and exit times for the exact law (3.4) and mean field model in units of the processing time T . Figure 2 shows the relative error in the processing times for the mean field model, i.e., the quantity

$$(3.8) \quad \left| \frac{e_n^{mf} - e_n^{exact}}{e_n^{exact} - a_n} \right|.$$

We see that the error in the processing times decreases as the number of parts in the system increases; i.e., (3.7) is a good approximation for large queues, as has to be expected of a mean field approximation.

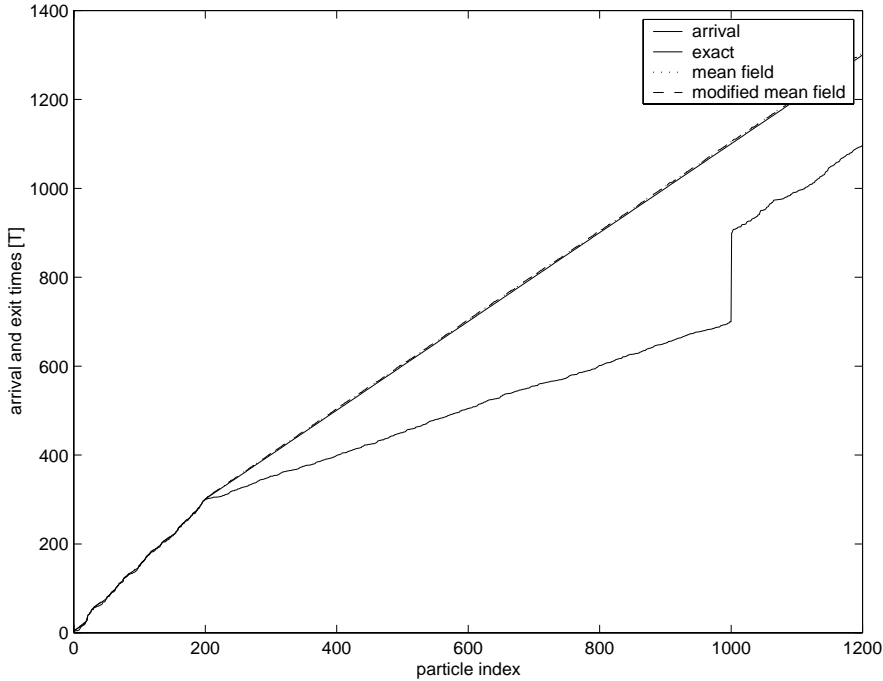


FIG. 1. Comparison between exit times for a simple queue computed with the exact model (3.4), the mean field model (3.6), and the modified mean field model in section 3.2.

Other than for the exact velocity model using the function ϑ_n^{ex} , we cannot guarantee that particle paths will not cross for the mean field problem (3.7). In other words, the solution of the mean field problem (3.7) will still consist of a superposition of δ -functions in x and u for given time t . However, for them to be supported at different values of x , they would have to satisfy, according to (3.2), the condition

$$(3.9) \quad \frac{1}{\vartheta^{mf}(a_{n-1}, \rho)} - \frac{1}{\vartheta^{mf}(a_n, \rho)} < a_n - a_{n-1} \quad \forall n.$$

If condition (3.2) is violated, this means that parts will not leave the system in the order they have arrived and $e_{n+1} < e_n$ holds for some n . Figure 3 shows the frequency with which this happens. The crossing of particle paths will have a major effect on the conservation law models in the next section, i.e., the occurrence of shocks in the Euler equations. One can avoid this problem by modifying the mean field model slightly and replacing (3.6) by

$$(3.10) \quad \vartheta^{mf}(t, \rho) = \frac{1}{T[1 + W(t-0) + \alpha(t)]}, \quad W(t) := \int_0^1 \rho(x, t) dx,$$

where the function $\alpha(t)$ is chosen such that (3.2) holds; i.e., we choose α such that

$$\frac{1}{\vartheta^{mf}(t, \rho)} - \frac{1}{\vartheta^{mf}(t+h, \rho)} \leq Kh \quad \forall h > 0$$

holds for some constant $0 < K < 1$. This means $\alpha(t)$ has to be chosen dependent

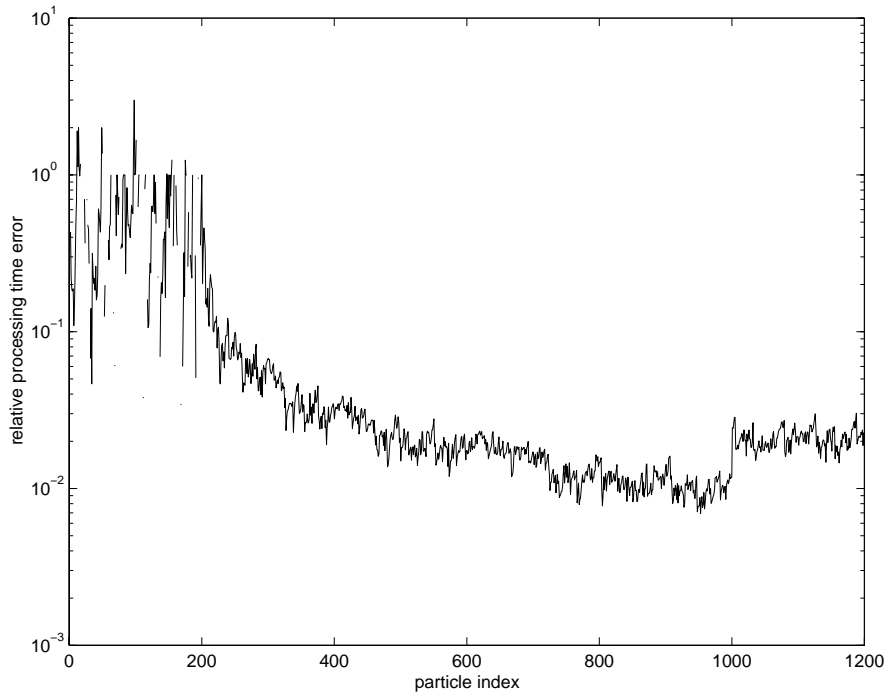


FIG. 2. Relative error in the processing times between the exact model and the mean field model for queues according to (3.8).

on $W(t)$ such that

$$\alpha(t+h) \geq \alpha(t) + W(t) - W(t+h) - \frac{Kh}{T}$$

holds for all $h > 0$. The use of the modified mean field model (3.10) guarantees that particle paths will not cross and that the conservation law solutions derived in the next section will not exhibit shocks (i.e., characteristics will never cross). It turns out that the use of this modification is of theoretical relevance only since we have never seen any discernible difference in numerical solutions computed with the mean field model and the modified mean field model. Figure 1 actually shows the exit times for the modified Little law as well.

3.3. Mean field models for supply chains. We close this section by outlining the approach in the case that the system described by the part model (1.1) and (1.2) is more complex than just a simple queue. A supply chain is a chain of suppliers that produce goods, both for one another and for a customer at the end of the chain. The flow through this chain is governed by the supply and demand between one member of the chain and the next (see [10] for an overview). If the degree of completion variable x now denotes the position of a part in the chain, then the principles formulated at the beginning of this section for a queue obviously do not hold anymore. Most importantly, values of x between zero and one now have a meaning, and therefore the velocities cannot be chosen as constants. We can, however, still formulate a mean field model along the lines of section 3.2. Let $T(x)$ denote the cumulative processing time of all the downstream suppliers. We still assume Little's law (3.5)(a) and write

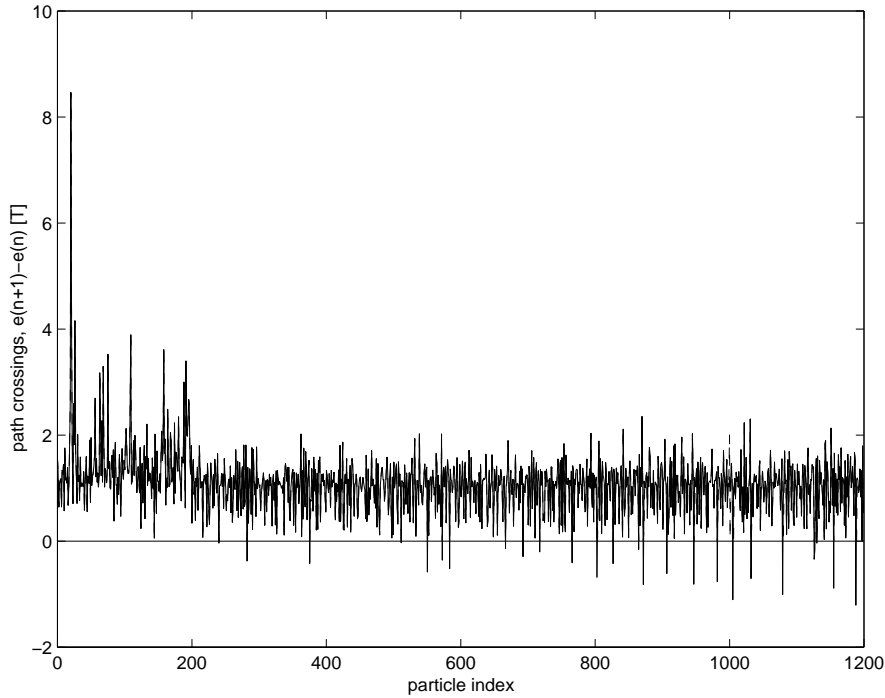


FIG. 3. Frequency of overtaking parts in the mean field model in section 3.2. $e_{n+1} - e_n < 0$ indicates overtaking.

$W = \tau\lambda$. However, W now denotes only the work in progress downstream in the chain, and τ denotes the time it takes the part to cover the distance from x to 1 in the chain. Other than in the queuing model, λ is now not some Markovian influx which is eliminated using (3.5)(b) but the flow function from one supplier to the next, i.e., the demand distribution in the chain, which is a given function determined by the policy of the supply chain. So Little's law now reads $W(x, t) = \tau(x, t)\lambda(x, t)$. We set the velocity ϑ now to

$$(3.11) \quad \vartheta(x, t, \rho) = \frac{1-x}{\tau(x, t)} = \frac{(1-x)\lambda(x, t)}{W(x, t)}, \quad W(x, t) := \int_x^1 \rho(y, t) dy,$$

meaning we estimate the velocity necessary to reach $x = 1$ at each point by the constant velocity which would be necessary if λ and W would not change in time. This is appropriate for a short period of time, after which the velocity estimate is updated together with the demand λ and the downstream load W . To model the effects of policies such as “pull” or “push” on the supply chain (see again [10] for a reference) the demand λ has to be made dependent on the density ρ itself.

Reentrant chains and critical loads. A reentrant supply chain is a chain in which a part has to pass through the same supplier, or workstation in a factory, more than once. This means that different values of x in our model denote the same physical station. It is observed experimentally that reentrant supply chains possess a critical workload. That means that, other than in (3.11), there is not a simple linear relation between the cycle time τ and the downstream workload W . In [2] and [3] a simple model for this problem has been analyzed, where the cycle time $\tau(W)$ has a pole at

$W = W^C$, the critical workload. The drawback of this model is that for values of W beyond W^C the resulting velocities become negative, and the model breaks down. In the context of mean field part models this can be expressed in the following way. The acceleration field S corresponding to (3.11) is given by

$$S(x, u, t) = u \frac{\partial}{\partial x} \frac{(1-x)\lambda}{W} + \frac{\partial}{\partial t} \frac{(1-x)\lambda}{W},$$

and we slow down the parts by adding a deceleration term which becomes large as the total work in progress $W(0, t)$ reaches the critical workload W^C , setting

$$S(x, u, t) = u \left[\frac{\partial}{\partial x} \frac{(1-x)\lambda}{W} - g(W^C - W(0, t)) \right] + \frac{\partial}{\partial t} \frac{(1-x)\lambda}{W}.$$

4. Conservation laws and continuum limits. As long as particle paths do not cross, which is always the case for the exact velocity law (3.4) and the modified law (3.10), and is true for Little's law (3.6) as long as condition (3.9) holds, the kinetic density $f(x, u, t)$ will be supported by a single δ -function in u for fixed (x, t) . Therefore, the velocity u can be eliminated as an independent variable, and the kinetic density f can be written as $f(x, u, t) = \rho(x, t)\delta(u - U(x, t))$. The equations for ρ and U , the compressible Euler equations, could be derived directly by insertion into the kinetic initial boundary value problem (3.7). If the maximal number of path crossings at any given point in space is known, one can replace the kinetic equation (2.2) by a system of moment equations whose order is this maximal number [14]. We would like to point out here an important difference to the otherwise very similar models for traffic flow. In gas dynamics models for traffic flows on multilane highways the overtaking of cars is a crucial component of the flow picture. Correspondingly, the interaction of cars with each other enters through a pressure term in the compressible Euler equations. The correct form of this pressure term has been the subject of debate (see [4], [9]), and the direct use of ideal gas laws leads to unrealistic results (i.e., negative velocities). Under the assumptions used in this paper the overtaking of parts is a rather rare event, caused by the inaccuracy of the mean field model for a linear queue. Therefore we will obtain the Euler equations at zero temperature. The interaction of different parts enters essentially through the boundary conditions, which contain all the important information about the mean field assumptions. This practical difference to traffic flow models arises from the fact that the "spatial" variable x here is an artificial construct; i.e., values of x between 0 and 1 have no special meaning.

Even if the parts in the system do not overtake each other and the paths do not cross, the resulting Euler equations would be posed with boundary data consisting of a superposition of δ -functions. From a computational point of view, this problem is actually not simpler than the solution of the Newton equations (1.3) for the individual parts. In this section we derive the limiting continuous problem for the moment equations for a large number of parts, under the assumption that the number of overtaking parts stays bounded. To be able to write down moment equations, it is necessary to express the moments of the acceleration term s in terms of the moments of the density function f . We therefore will assume for the rest of this section that the acceleration field is polynomial in the velocity u , i.e., that

$$S(x, u, t) = \sum_{j=0}^J S_j(x, t) u^j$$

holds. We consider a finite number N of parts arriving in the system in the time interval $[0, T]$; i.e., $0 = a_1 < \dots < a_N = \hat{T}$ holds. In order to carry out the continuum limit, we first have to bring the kinetic problem (2.2) into an appropriate dimensionless form. We employ the scaling

$$t = \hat{T}t_s, \quad u = \frac{u_s}{\hat{T}}, \quad f = \hat{T}Nf_s, \quad a_n = \hat{T}a_{ns}, \quad S = \frac{S_s}{\hat{T}^2}, \quad \vartheta^0 = \frac{1}{\hat{T}}\vartheta_s^0, \quad v_n = \frac{v_{ns}}{\hat{T}},$$

and obtain the scaled version of the kinetic problem (2.2):

(4.1)

$$\partial_{t_s} f_s + u_s \partial_x f_s + \partial_{u_s} (S_s f_s) = 0, \quad u_s f_s(0, u_s, t_s) = \frac{1}{N} \delta(u_s - \vartheta_s^0(t_s, A, f_s)) \sum_{n=1}^N \delta(t_s - a_{ns}).$$

We will drop the subscript s again from now on. In order to arrive at a continuous problem we smooth out the boundary data. To this end we choose a convolution kernel $G(t)$, which is compactly supported in $[0, 1]$ and satisfies $\int G(t) dt = 1$, and replace the superposition of δ -functions in (4.1) by

$$(4.2) \quad \lambda(t) = \frac{1}{N\varepsilon} \int G\left(\frac{t-\tau}{\varepsilon}\right) \sum_{n=1}^N \delta(\tau - a_n) d\tau = \frac{1}{N\varepsilon} \sum_{n=1}^N G\left(\frac{t-a_n}{\varepsilon}\right),$$

where ε denotes a small, dimensionless parameter. We have the following theorem.

THEOREM 2. *Let the number N of parts in the system be large. Let the number of parts passing other parts be bounded by $M \ll N$. Furthermore, let the velocities v_n be uniformly bounded for all n ; then the density $\rho(x, t) := \int f(x, u, t) du$ satisfies the initial boundary value problem*

(4.3)

$$(a) \quad \partial_t \rho + \partial_x (\rho U) = 0, \quad \partial_t (\rho U) + \partial_x \left(\rho U^2 + \frac{M}{N} \omega_2 \right) - \sum_{j=0}^J S_j \left(\rho U^j + \frac{M}{N} \omega_j \right) = 0,$$

$$(b) \quad \rho U(0, t) = \lambda(t) + \varepsilon R(t), \quad \left(\rho U^2 + \frac{M}{N} \omega_2 \right) (0, t) = \vartheta(t, 0, \rho) (\lambda(t) + \varepsilon R(t)),$$

weakly in space and time, where the functions ω_j are bounded in \mathcal{L}^1 and λ is defined by (4.2). The remainder term R in the boundary conditions is bounded in \mathcal{W}_1^∞ .

Proof of Theorem 2. We define the moments $m_k(x, t)$ of the kinetic density function f via

$$(4.4) \quad m_k(x, t) = \int u^k f(x, u, t) du = \frac{1}{N} \sum_{n=1}^N v_n^k \delta(x - z_n(t)) H(t - a_n).$$

Because of the kinetic problem (3.7) the moments m_k satisfy the system

$$(4.5) \quad (a) \quad \partial_t m_k + \partial_x m_{k+1} - k \sum_{j=0}^J S_j m_{k+j-1} = 0,$$

$$(b) \quad m_k(x, 0) = 0, \quad m_{k+1}(0, t) = \frac{1}{N} \vartheta(t, 0, m_0)^k \sum_{n=1}^N \delta(t - a_n), \quad k = 0, 1, \dots$$

We split the population of parts into two groups, namely those which do not overtake as long as they are inside the interval $[0, 1]$ and the rest. We denote the indices corresponding to these two groups with \mathcal{N}_1 and \mathcal{N}_2 . So $\{1, \dots, N\} = \mathcal{N}_1 \cup \mathcal{N}_2$ and $|\mathcal{N}_2| = M$ holds. Let $U(x, t)$ be a smooth function which interpolates the values of $v_n(t)$ at $z_n(t)$, i.e.,

$$U(z_n(t), t) = v_n(t), \quad n \in \mathcal{N}_1, \quad 0 \leq z_n(t) \leq 1.$$

The function U is well defined since, by definition, the values of $z_n(t)$, $n \in \mathcal{N}_1$, are distinct for every given t as long as they are within the interval $[0, 1]$. We now write the moments m_k as

$$m_k(x, t) = m_0 U^k + \frac{M}{N} \omega_k, \quad k = 1, \dots.$$

Using the definition (4.4) of the moments m_k gives

$$\omega_1 = 0, \quad \omega_k(x, t) = \frac{1}{M} \sum_{n \in \mathcal{N}_2} \delta(x - z_n(t)) H(t - a_n) (v_n(t)^k - U(z_n(t), t)^k), \quad k \geq 2,$$

and the initial boundary value problem for the zero and first order moment equations in (4.5) becomes

$$(4.6) \quad (\text{a}) \quad \partial_t m_0 + \partial_x(m_0 U) = 0,$$

$$\partial_t(m_0 U) + \partial_x \left(m_0 U^2 + \frac{M}{N} \omega_2 \right) - \sum_{j=0}^J s_j \left(m_0 U^j + \frac{M}{N} \omega_j \right) = 0,$$

$$(b) \quad m_0 U(0, t) = \alpha(t) := \frac{1}{N} \sum_{n=1}^N \delta(t - a_n), \quad \left(m_0 U^2 + \frac{M}{N} \omega_2 \right) (0, t) = \vartheta(t, 0, N m_0) \alpha(t).$$

After the obvious change of notation from m_0 to ρ , these are the equations (4.3), except for the boundary data in (4.6)(b). The boundary datum α is a measure, and (4.6) still has to be understood in the weak sense. Computing the difference between α and λ , defined in (4.2), when integrated against a test function ψ gives

$$\begin{aligned} \int \psi(t) (\lambda(t) - \alpha(t)) dt &= \frac{1}{N} \sum_{n=1}^N \left[\int \frac{1}{\varepsilon} \psi(t) G \left(\frac{t - a_n}{\varepsilon} \right) dt - \psi(a_n) \right] \\ &= \frac{1}{N} \sum_{n=1}^N \int [\psi(a_n + \varepsilon t) - \psi(a_n)] G(t) dt = \frac{\varepsilon}{N} \sum_{n=1}^N \int t \psi'(a_n + \varepsilon \xi(t)) G(t) dt, \end{aligned}$$

giving for the remainder R in (4.3)(b)

$$\left| \int \psi(t) R(t) dt \right| \leq \text{const} \|\psi'\|_\infty. \quad \square$$

So, under the assumption that $M \ll N$ holds, we will, after choosing a value for the smoothing parameter ε , neglect the closure functions ω_j and the residual R in (4.3) and solve a version of the compressible Euler equations at zero temperature with smoothed boundary data.

Remark. If the overtaking of parts is a significant occurrence, i.e., the assumption $M \ll N$ is not justified, we are faced with the usual closure problem of defining an equation of state for the “temperature,” the variance in the velocities of parts sharing the same point in space. This can be expected to be the case for reentrant supply chain models in section 3.3 but not for models of simple linear chains.

Remark. We should comment here on the role of the smoothing parameter ε . On one hand, ε should be small to allow us to neglect the remainder term R in (4.3)(b). On the other hand, ε has to be large enough such that the smoothed influx density λ is still a smoothly varying function. Computing the derivative of λ in (4.2) gives that $\lambda'(t)$ is roughly of the order $\frac{1}{N\varepsilon^2}$ times the number of parts arriving in the time interval $[t, t + \varepsilon]$. If the parts would arrive at a uniform rate λ_0 , then the size of λ' would be bounded by to $\frac{\lambda_0}{N\varepsilon}$, and we should set $\frac{1}{N} \ll \varepsilon \ll 1$, in order to obtain a constant λ . Making ε smaller reduces the error in (4.3)(b), i.e., allows for a better resolution of the fluctuations in the arrival times, at the expense of having to resolve more structured boundary data.

Remark. In the case that the velocity functions $\vartheta_n(t, A, Z(\tau), V(\tau), \tau \leq t)$ are given explicitly in mean field form, that is, if model ϑ_n directly and not implicitly via the s_n , (4.3) reduces to a scalar, almost linear, conservation law for the density ρ . If the ϑ_n are of the form $\vartheta_n = \vartheta(t, z_n(t), \rho)$, then parts cannot overtake since all parts sharing the same point in space will have the same velocity. Therefore, the ω_j in (4.3) will vanish. On the other hand, the acceleration field S will be of the form

$$S = \partial_t \vartheta + u \partial_x \vartheta + D_\rho \vartheta [\partial_t \rho],$$

where $D\vartheta$ denotes the functional derivative with respect to ρ . This gives the equation

$$\partial_t(\rho U) + \partial_x(\rho U^2) - (\partial_t \vartheta + D_\rho \vartheta [\partial_t \rho])\rho - \rho U \partial_x \vartheta = 0,$$

which, using the conservation law for ρ , has a solution $U(x, t) = \vartheta(t, x, \rho)$. Of course, in this case the conservation law for the density ρ could have been derived directly, and the detour over the kinetic problem and the moment closures is unnecessary.

5. Numerical results. In this section we verify the mean field model (3.6) together with the continuum approximation and the moment closure as given by Theorem 2 in the previous sections numerically for the case of a simple queue. Combining the mean field part velocity model (3.6) from section 3.2 with the moment closure yields the compressible Euler equations at zero temperature, i.e.,

$$(5.1) \quad \partial_t \rho + \partial_x(\rho U) = 0, \quad \partial_t(\rho U) + \partial_x(\rho U^2) = 0.$$

There are no acceleration terms since each part moves with constant velocity. The mean field model enters only through the boundary conditions, which are of the form

$$(5.2) \quad \rho U(0, t) = \lambda(t), \quad \rho U^2(0, t) = \vartheta(t, 0, \rho) \lambda(t) = \frac{\lambda(t)}{T[1 + W(t)]}, \quad W(t) = \int_0^1 \rho(x, t) dx,$$

where the function $\lambda(t)$ is the smoothed out superposition of δ -functions given by (4.2).

As a test example we choose the same set of randomly chosen arrival times as in section 2. We choose 1100 random arrival times, where the first 200 parts arrive at a rate of 0.666 parts per processing time T and the next 800 parts arrive at a rate of 2 parts per processing time T . Nothing arrives for the next 200 time units, and

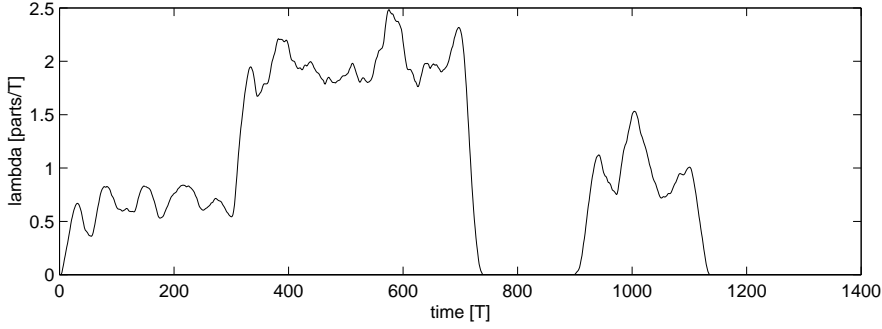


FIG. 4. Smoothed out influx rate λ as given by (4.2).

then another 200 parts arrive at a rate of 1 part per processing time. We compare the numerical solution of the Euler equations (5.1) to the solution of the exact part model, i.e., using the sequence of velocities generated by ϑ_n^{ex} in (3.4), and the mean field part model using (3.6). We use a value of $\varepsilon = 0.01$ for the smoothing parameter ε in (4.2). Using $O(10^3)$ parts, this means that we resolve the δ -functions quite well at the expense of resolving a rather rough influx function λ . For the smoothing function G in (4.2) we choose

$$G(\tau) = \begin{cases} 6\tau(1-\tau) & \text{for } 0 \leq \tau \leq 1 \\ 0 & \text{else} \end{cases}.$$

Figure 4 shows the corresponding smoothed influx function $\lambda(t)$ in (4.2). We use the simplest possible method for the solution of the smoothed conservation law problem, that is, Godunov's method with piecewise constant reconstruction (see [15]). Since the velocities U will always remain nonnegative, Godunov's method reduces to simple upwinding. Defining the quantities

$$(\rho U^s)_{jk} := \frac{1}{\Delta x} \int_{x_j}^{x_{j+1}} \rho(x, t_k) U(x, t_k)^s dx, \quad F_{jk}^s = \frac{1}{\Delta t_k} \int_{t_k}^{t_{k+1}} \rho(x_j, t) U(x_j, t)^{s+1} dt,$$

for an equidistant mesh in x -direction $x_j = j\Delta x$, $j = 0, \dots, J$, $J\Delta x = 1$, and a variable time step Δt_k , we obtain

$$(5.3) \quad (\rho U^s)_{j,k+1} = (\rho U^s)_{j,k} - \frac{\Delta t_k}{\Delta x} [F_{j+1,k}^s - F_{jk}^s], \quad j = 0, \dots, J-1, \quad s = 0, 1.$$

Godunov's method for the present problem now consists simply of setting $F_{jk}^s = (\rho U^{s+1})_{j-1,k}$ for $j = 1, \dots, J-1$ and $s = 0, 1$. The fluxes F_{0k}^s , $s = 0, 1$ are given by the boundary condition (5.2), i.e.,

$$(5.4) \quad F_{0k}^0 = \frac{1}{\Delta t_k} \int_{t_k}^{t_{k+1}} \lambda(t) dt, \quad F_{0k}^1 = \frac{1}{\Delta t_k} \Phi[\rho, U](t_k) \int_{t_k}^{t_{k+1}} \lambda(t) dt,$$

where the velocity model Φ is evaluated at the previous time step. Figures 5 and 6 show the density ρ and the velocity U as a function of x and the time t in units of the processing time T on a logarithmic scale. Figure 7 shows the total work in progress, i.e., the quantity

$$(5.5) \quad W(t) = \int_0^1 \tilde{\rho}(x, t) dx$$

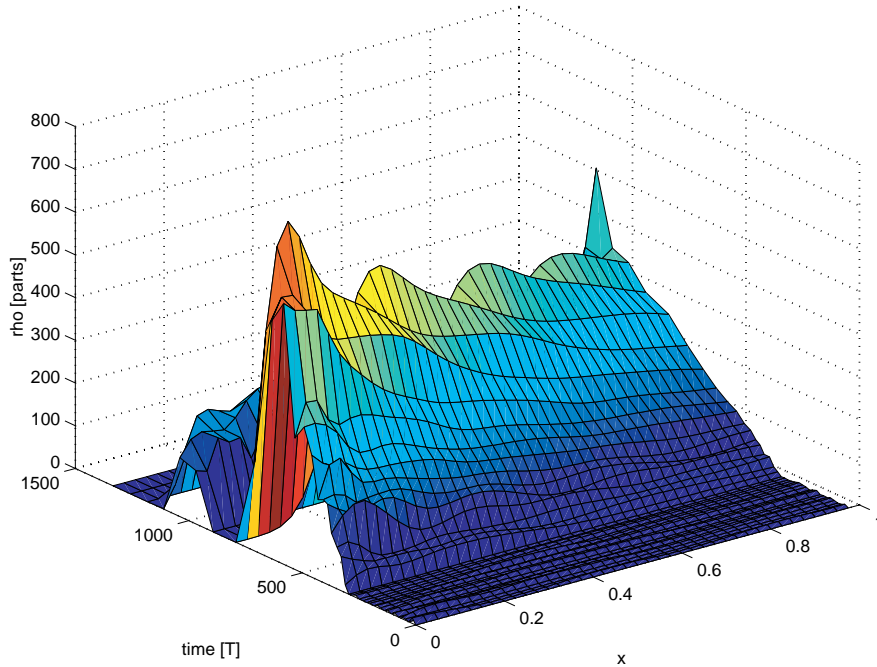


FIG. 5. Part density ρ computed from the Euler equations (5.1).

as a function of time. The total work in progress is computed from the numerical solution of the conservation law (5.3) and from the smoothed part model solutions using the sequences generated by ϑ_n^{ex} and ϑ^{mf} in (3.4) and (3.6). So we compare (5.5) to the quantities

$$W^{part}(t) = \int_0^1 dx \int_0^\infty d\tau G(\tau) \rho^{part}(x, t - \varepsilon\tau), \quad \rho^{part}(x, t) = \sum_{n=1}^N \delta\left(x - \frac{t - a_n}{e_n - a_n}\right) H(t - a_n),$$

where the exit times e_n are computed using ϑ_n^{ex} and ϑ^{mf} . We see that the largest error occurs around $t = 1000T$, the time when the queue runs precisely on capacity and the expected work in progress should remain constant in time. This is expected, since in this case all the velocities predicted by the mean field model should be the same, and the statistical fluctuations in the arrival times lead to overtaking of parts. Figure 8 shows snapshots of the comparison between the density ρ and the corresponding density computed from the particle model using the corrected Little law (3.10), i.e., the quantity

$$\rho_j^{part}(t) = \int_{x_j}^{x_{j+1}} dx \int_0^\infty d\tau G(\tau) \rho^{part}(x, t - \varepsilon\tau).$$

Remark. Other than in discrete part models, the computational effort of solving the initial boundary value problem (5.1), (5.2) for the Euler equations actually decreases with an increasing number of parts in the system. An increasing number of parts, i.e., larger values of the work in progress function W in (5.2), translates into smaller velocities, allowing for larger time steps still satisfying the CFL condition of the Godunov method. This reflects the fact that for system with a large number of

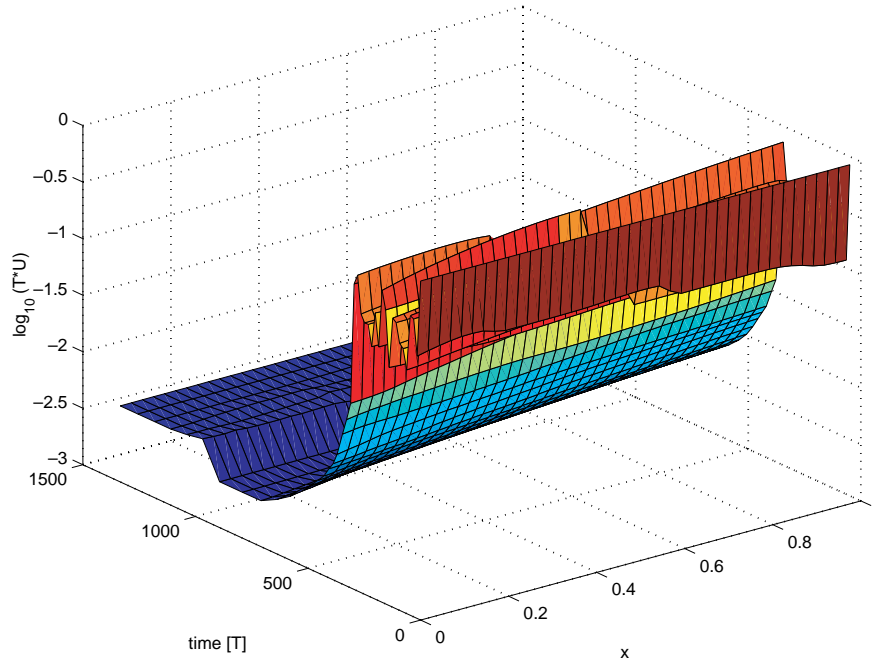


FIG. 6. Velocity U computed from the Euler equations (5.1) (logarithmic scale).

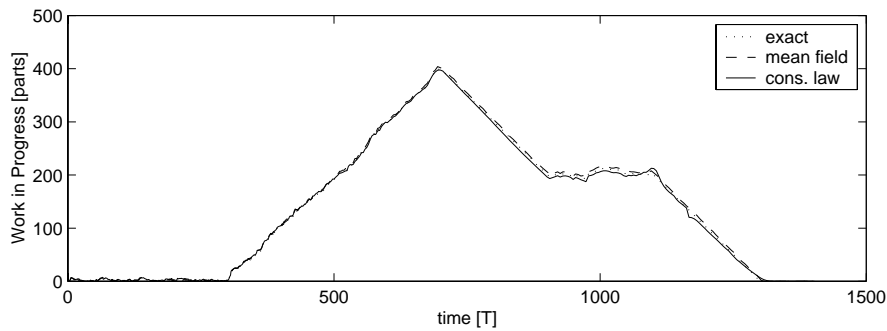


FIG. 7. Work in progress W computed from the conservation law, the exact part model, and the mean field part model in (5.5) as a function of time.

parts the scaled variance of their velocities will decrease, and they can therefore be “lumped” in the discrete integral influx boundary condition (5.4).

Remark. The system (5.1) does, of course, admit shock solutions. However, since it is derived, according to Theorem 2, as a small perturbation of a system where characteristics do not intersect, these shocks will be negligibly small as long as the mean field boundary condition (5.2) is a good approximation of the exact picture.

6. Conclusions. In this paper we have presented a framework for connecting supply chain models based on the motion of individual parts in the system to models for the integral density of these parts, consisting essentially of the equations of gas dynamics. These gas dynamic models represent a compromise between discrete

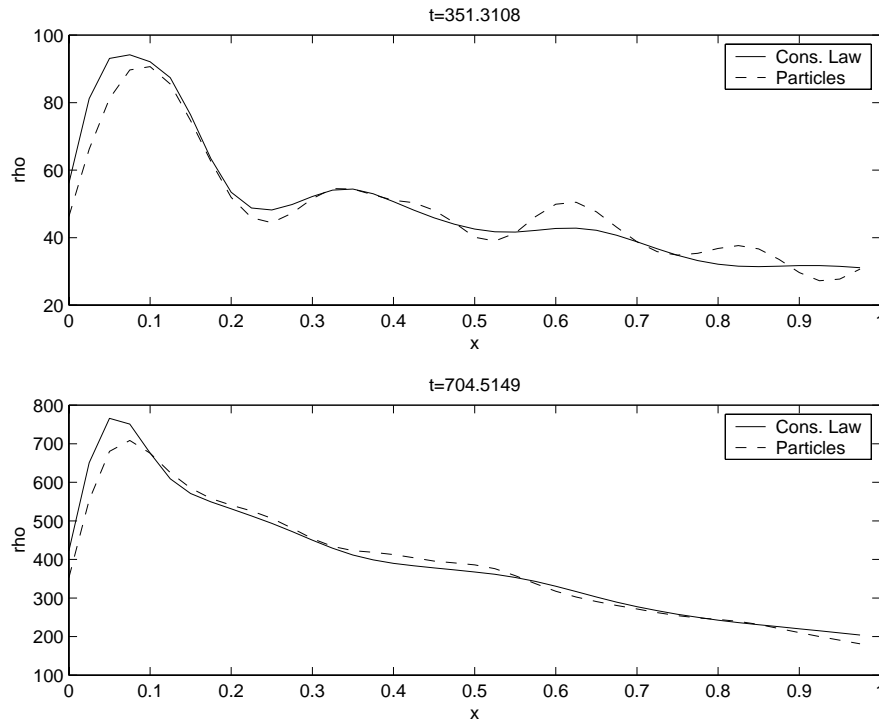


FIG. 8. Comparison of conservation law density and exact part model density ρ at $t \approx 350T$ and $t \approx 700T$.

event simulators and rate equation models (so-called fluid models) in that they are computationally more efficient than discrete event simulations and provide more information and are more realistic than rate equations. The increase in efficiency is due to the fact that parts are still modeled as a continuum; i.e., the computational cost of gas dynamic models does not increase with the number of parts considered. These models bear strong similarities to models for traffic flow on highways for which similar kinetic and gas dynamic hierarchies can be derived. The basic difference to traffic flow models arises from the fact that nearest neighbor interactions play a much more significant role in traffic flow, while mean field theories for supply chains have to be derived from global statistical properties.

In the special case of simple queues, the individual part models can be formulated explicitly, and we have verified the validity of the fluid approximations numerically. For models of more complex supply chains (especially reentrant chains) the motion of parts has to be modeled by more phenomenological approaches, which, however, render themselves to the same model hierarchy approach.

REFERENCES

- [1] E. J. ANDERSON, *A new continuous model for job shop scheduling*, *Internat. J. Systems Sci.*, 12 (1981), pp. 1469–1475.
- [2] D. ARMBRUSTER, D. MARTHALER, AND C. RINGHOFER, *Modeling a re-entrant factory*, *Oper. Res.*, submitted.
- [3] D. ARMBRUSTER, D. MARTHALER, AND C. RINGHOFER, *A mesoscopic approach to the simulation of semiconductor supply chains*, in *Proceedings of the International Conference on*

- Modeling and Analysis of Semiconductor Manufacturing (MASM2002), G. Mackulak, J. Fowler, and A. Schömig, eds., Tempe, AZ, 2002, pp. 365–369.
- [4] A. AW AND M. RASCLE, *Resurrection of “second order” models of traffic flow*, SIAM J. Appl. Math., 60 (2000), pp. 916–938.
 - [5] J. BANKS, J. CARSON, II, AND B. NELSON, *Discrete Event System Simulation*, Prentice–Hall, Upper Saddle River, NJ, 1999.
 - [6] R. BILLINGS AND J. HASENBEIN, *Applications of fluid models to semiconductor fab operations*, in Proceedings of the 2001 International Conference on Semiconductor Manufacturing Operational Modeling and Simulations, Seattle, WA, 2001.
 - [7] P. BRANDT, A. FRANKEN, AND B. LISEK, *Stationary Stochastic Models*, John Wiley, New York, 1990.
 - [8] C. CERCIGNANI, *The Boltzmann Equation and Its Applications*, Appl. Math. Sci. 67, Springer-Verlag, New York, 1988.
 - [9] C. DAGANZO, *Requiem for second order fluid approximations of traffic flow*, Transportation Res. Part B, 29 (1995), pp. 277–286
 - [10] C. DAGANZO, *A Theory of Supply Chains*, ITS Research report UCB-ITS-RR2001-7, <http://www.ce.berkeley.edu/~daganzo/> (2001).
 - [11] M. EL-TAHA AND S. STIDHAM, *Sample Path Analysis of Queuing Systems*, Internat. Ser. Oper. Res. Management Sci. 11, Kluwer Academic Publishers, Boston, 1999.
 - [12] D. HELBING, *Gas kinetic derivation of Navier Stokes like traffic equations*, Phys. Rev. E (3), 53 (1996), pp. 2366–2381.
 - [13] D. HELBING, *Traffic and related self-driven many particle systems*, Rev. Modern Phys., 73 (2001), pp. 1067–1141.
 - [14] S. JIN AND X. T. LI, *Multi-phase Computations of the Semiclassical Limit of the Schrodinger Equation and Related Problems: Whitham vs. Wigner*, preprint, <http://www.math.wisc.edu/~jin/research.html> (2002).
 - [15] R. LEVEQUE, *Numerical Methods for Conservation Laws*, Birkhäuser, Basel, 1992.
 - [16] M. LIDTHILL AND J. WHITHAM, *On kinematic waves. I. Flood movement in long rivers*, Proc. Roy. Soc. London Ser. A, 229 (1955), pp. 281–316.
 - [17] M. LIDTHILL AND J. WHITHAM, *On kinematic waves. II. A theory of traffic flow on long crowded roads*, Proc. Roy. Soc. London Ser. A, 229 (1955), pp. 317–345.
 - [18] I. PRIGOGINE AND R. HERMAN, *Kinetic Theory of Vehicular Traffic*, Elsevier, New York, 1971.
 - [19] M. C. PULLAN, *An algorithm for a class of continuous linear programs*, SIAM J. Control Optim., 31 (1993), pp. 1558–1577.