Analysis of Airplane Boarding Times

Eitan Bachmat [*] Daniel Berend [†] Luba Sapir [‡] Steven Skiena [§]
Natan Stolyarov [¶]

Abstract

We model and analyze the process of passengers boarding an airplane. We show how the model yields closed-form estimates for the expected boarding time in many cases of interest. Comparison of our computations with previous work, based on discrete event simulations, shows a high degree of agreement. Analysis of the model reveals a clear link between the efficiency of various airline boarding policies and a congestion parameter which is related to interior airplane design parameters, such as distance between rows. In particular, as congestion increases, random boarding becomes more attractive among row based policies.

1 Introduction

The process of airplane boarding is experienced daily by millions of passengers worldwide. Airlines have adopted a variety of boarding strategies in the hope of reducing the gate turnaround time for airplanes. Significant reductions in gate delays would improve on the quality of life for long-suffering air travelers, and yield significant economic benefits from more efficient use of aircraft and airport infrastructure. (See Van Landeghem and Beuselinck (2002), Marelli *et al.* (1998) and Van den Briel *et al.* (2005).)

The most pervasive strategy currently employed links passenger queueing time to row assignment. In particular, airlines tend to board passengers from the back of the airplane first. Such "back-to-front" policies are implemented by announcements of the form "Passengers from rows 30 and above are now welcome to board the plane".

It is not clear *a priori* how to analyze such strategies or to determine which policies are most effective at minimizing the expected boarding time. Airplane boarding has been previously studied through discrete event simulations by Van Landeghem and Beuselinck

[*]Department of Computer Science, Ben-Gurion University, Email: ebachmat@cs.bgu.ac.il, Research partially supported by an IBM faculty award

[†]Departments of Mathematics and Computer Science, Ben-Gurion University, Email: berend@cs.bgu.ac.il

[‡]Department of Industrial Engineering and Management, Ben-Gurion University, Email: lsapir@bgumail.bgu.ac.il

[§]Department of Computer Science, SUNY at Stony Brook, Email: skiena@cs.sunysb.edu

[¶]Department of Computer Science, Ben-Gurion University, Email: stolyarov@cs.bgu.ac.il

(2002), Marelli *et al.* (1998) and Ferrari and Nagel (2005). In addition, Van den Briel *et al.* (2003) formulated a non-linear integer programming problem, related to airplane boarding time, to which they applied various heuristics in order to find efficient boarding policies. The policies were then tested using a discrete event simulation.

Somewhat surprisingly, these studies have found that back-to-front policies are not necessarily effective, and might even be detrimental when compared with a boarding process with no boarding policy at all (random boarding). Van Landeghem and Beuselinck (2002) argue that back-to-front policies are ineffective because they cause local congestion in the airplane, but no explanation is given for the mechanism by which congestion affects boarding time.

These studies also show that outside-in boarding policies, in which window seat passengers board first, followed by middle and then aisle seat passengers, can improve boarding time. In such policies, and others suggested by the studies, passengers are divided into multiple classes or groups which are boarded in sequence.

The study by Marelli *et al.* (1998) of Boeing Corp. emphasizes the effect of airplane interior design on boarding time, again using discrete event simulation methods. The paper describes a commercial simulation product, and simulation results are only sketched. It is therefore difficult to analyze the results of this work.

The results and observations of Van Landeghem and Beuselinck (2002), Van den Briel *et al.* (2003) and Ferrari and Nagel (2005) are of considerable value and interest. However, they do not address the need for a unified, analytic approach which can lead to a deeper understanding of the boarding process. As an example of the limitations of previous methods, we note that the simulations in all studies were carried out with particular airplanes in mind. It is therefore important to know how the success of airline boarding policies is related, if at all, to airplane design parameters, such as distance between rows, if the results are to be extended to other airplanes. The model of Van den Briel *et al.* (2003) does not take such design parameters into account.

In addition, any new scenarios need to be tested from scratch, and there is very little insight to be gained as to the underlying mechanisms of success for various strategies.

An analytical model which takes airplane design parameters into account was recently introduces in Bachmat *et al.* (2005, 2006). The authors express the expected boarding time T, as the number of passengers is large, in terms of the solution to a variational problem. The variational problem turns out to have a beautiful geometric interpretation in terms of spacetime (Lorentzian) geometry. A major issue that is not addressed in Bachmat *et al.* (2005, 2006) is the validation of the model against detailed simulations of the boarding process. This is particularly important since the model makes many simplifying assumptions. In addition, the variational problem which is solved describes the asymptotic behavior of the model when the number of passengers tends to infinity. The papers also use a substantial amount of physics-related terminology, a fact which makes them less accessible to the operations research community.

The purpose of this paper is to:

- explain the model in terms of classical notions in operations research and optimization;
- explain how to use the model to compute boarding times and assess boarding policies.;
- validate the model against results from detailed computer simulations;
- establish the relation between congestion and efficiency of airline boarding policies. When congestion is low, back-to-front boarding policies are very effective, while random boarding is very ineffective. When congestion is high, random boarding is very effective, while back-to-front boarding with equal size groups is detrimental.

In establishing the relation between congestion and airline boarding policies we utilize a clear advantage of analytical models over simulations, namely the ability to consider large classes of boarding policies at once. In Section 2 we consider airplane boarding as a project consisting of tasks with precedence relations. When passengers board the airplane they essentially "solve" the project scheduling problem via the critical path method. The input parameters for the project, which determine the precedence relations and delay times, include distance between rows in the airplane, number of passengers per row, passenger aisle clearing time and the airline boarding policy. Airplane boarding is not a deterministic process but rather a stochastic one. We explain how to provide a stochastic framework to the input parameters and thus to the boarding process.

As is typical of many stochastic processes, the behavior of the process becomes simpler and more deterministic as the number of passengers tends to infinity. In Section 3 we present, following Bachmat *et al.* (2005, 2006), a constrained variational problem whose solution provides the expected critical path and boarding time in the limit as the number of passengers tends to infinity. The input to the variational problem is a probability measure p, determined by the airline boarding policy and a congestion parameter k which is related to the interior design of the airplane.

In Section 4 we describe a large class of airline boarding policies. We show how to associate with to each policy a probability measure p, to be used in the variational problem.

In Section 5 we provide detailed computations of the expected boarding times for a large class of airline boarding policies. The computations show how the variational problem can be solved in many cases and form the basis for comparison with the simulation results.

In Section 6 we apply the calculations of Section 5 to many specific boarding policies considered in previous studies. A comparison of our calculations with the results of the detailed discrete event simulations of Van Landeghem and Beuselinck (2002) and Van den Briel *et al.* (2003) shows a rather remarkable degree of overall agreement on the relative merit of various policies. The comparison serves to validate the model. In addition, the results show that it is difficult to substantially improve upon random boarding (i.e., no policy) without exerting a large degree of control over the boarding process or reverting to policies which incorporate outside-in seat boarding. It seems unlikely that passengers will tolerate strict control boarding policies, even if these are easy to implement. Among policies which exert mild control over passengers, the best candidates for improving upon random boarding are policies combining outside-in boarding with a touch of back-to-front boarding.

In Section 7 we prove a general result which holds for a very large class of boarding policies. The result shows that the failure of Van Landeghem and Beuselinck to find mildly intrusive, row based, boarding policies, which will improve upon random boarding, is not coincidental. In fact, random boarding is asymptotically optimal as congestion grows among all row based policies, not only those studied in the simulations.

Section 8 provides a brief summary and suggestions for future work.

2 The airplane boarding process

In this section we provide a model for the airplane boarding process. Our approach is to think of airplane boarding as a project and use the standard representation of a project as a network with precedences. We view airplane boarding as consisting of tasks, one per passenger. The task of each passenger is to sit down. The precedence relations come from aisle blocking, where a given passenger blocks the aisle while handling his/her luggage and getting organized before sitting down. While this activity is taking place, other passengers which are behind the given passenger cannot proceed to their assigned rows. Once the passenger's task is complete, they clear the aisle. As an example, consider passenger A whose row is 25 and passenger B who is further back in the boarding queue and sits at row 30. It is obvious that, until passenger A clears the aisle, passenger B will not be able to reach his/her row. The delay associated with a passenger is the total amount of time that the passenger blocks the aisle once he has reached his designated row.

Once we consider airplane boarding as a network of tasks with execution times (the delays) and precedence relations, we notice that the boarding process essentially coincides with the critical path method (CPM).

We formalize the network/project approach as follows. We assume that passengers are assigned seats in the airplane in advance of the boarding process. Boarding is from the front of the plane, and the front row is row 1.

The input data which determines the network is composed of the following items.

- A number h which represents the number of passengers per row.
- A sequence of passengers $x_1, \ldots, x_n$, where x_i denotes the i'th passenger in the boarding queue. Each passenger x_i has a seat in an assigned row, denoted by $r_i = r(x_i)$. The row number r_i ranges between 1 and n/h, and each number in that range is the row number of h different passengers. We implicitly assume here that the airplane is full.

- A length value W, which measures the aisle length occupied by a passenger, baggage and personal space included.

- A delay value D, which measures the total amount of time it takes from the moment passenger x_i has reached his/her designated row until he clears the aisle. This time includes getting organized, placing carry-on luggage and, possibly, passing by previously seated passengers from the same row on the way to the designated seat. The last operation usually requires the seated passengers to get up and sit back after the newly arrived passenger has taken his/her seat.

- A length parameter l, representing the distance between successive rows.

The delay associated with the i'th passenger/task is D. The precedence relations are defined as follows:

The location of row m along the aisle is ml. When passenger x_i reaches his/her row, r_i, he/she occupies the aisle length segment from lr_i to $lr_i + W$. Passengers arriving at time t_i to their assigned row sit down and clear the aisle at time $t_i + D$. Once they clear the aisle, passengers who are behind them continue marching along the aisle as far as they can, and the process repeats.

In terms of these parameters, we define the precedence/blocking relations as follows. Consider the time $t = t_i + D - \varepsilon$, just before the i'th passenger clears the aisle. Let $x_{i_1}, \ldots, x_{i_k}$ be the passengers who are lined up behind x_i at time t. Passenger x_i *has precedence over* passenger x_{i_j} if

$$lr_i - (j-1)W < lr_j. \tag{1}$$

The precedence relation can be explained as follows. Since $x_{i_1}, \ldots, x_{i_{j-1}}$ are lined up behind x_i, they create an aisle backlog which stretches back to at least $lr_i - (j-1)W$, and if the inequality holds, this prevents x_j from getting to his/her row.

It is customary, when dealing with projects, to represent the project using a 2-dimensional chart, in which tasks are represented by points and precedence relations are represented by arrows pointing to the right. We represent the airplane boarding process as follows. Let $m = n/h$ be the number of rows. Consider the passenger x_i, who is in the i'th position in the boarding queue and sits at row r_i. We represent x_i by the point $(i/n, r_i/m)$ in the unit square. Since passengers can only block passengers who are further than them in the queue, precedence relations will be represented by arrows pointing to the right.

Airplane boarding is not a deterministic process since airlines do not have full control on the order in which passengers queue. We therefore need to provide a stochastic process setting for the boarding process. We achieve this goal by presenting the ordering of passengers x_i in a random variable setting, in which the representation $(i/n, r_i/m)$ is relaxed.

Let $p(q, r)$ be a joint distribution function on the unit square $0 \leq q \leq 1$, $0 \leq r \leq 1$, which represents passengers' row and queue positions. We will show later on how $p(q, r)$ is determined by the airline's boarding policy (and also by how passengers react to this policy). It is through p that we model and compare boarding policies.

To determine the ordering of passengers, the distribution $p(q,r)dqdr$ is sampled n times independently to produce passenger coordinates $x_i = (q_i, \tilde{r}_i)$, $i = 1, \ldots, n$, where we assume that passengers are indexed in increasing order of the q-coordinate, i.e., $q_1 < q_2 < \ldots < q_n$. To determine the row $r_i = r(x_i)$ of passenger x_i, the passengers are sorted by the value of $\tilde{r}_i$ in increasing order. The first h passengers are assigned to seats in row 1, the next h to seats in row 2 and so on.

We are interested in studying the vector (t_i), where t_i is the time passenger x_i takes his seat. In particular, we are interested in the behavior of the random variable $t = \max_i \ t_i$, which represents the project completion time (length of critical path), or total boarding time.

3 Modeling the asymptotic behavior of airplane boarding with space-time

Consider the stochastic airplane boarding process with parameters n, D, h, l, W, p, as defined in the previous section. Many probabilistic systems obey laws of large numbers, which makes their behavior more predictable as the number of system components increases. We will therefore consider the asymptotic behavior of the boarding time random variable t as the number of passengers tends to infinity. We will eventually have to verify that the results we obtain are relevant for finite populations as well. Put

$$\alpha(q,r) = \int_r^1 p(q,z)dz \tag{2}$$

and

$$k = hW/l. \tag{3}$$

We make the technical assumption that the unit square may be divided into a finite number of regions in such a way that p is differentiable in each region (p is piecewise differentiable). This assumption will hold in all our examples.

Let $T = T(D,h,l,W,p) = T(p,k)$ be the solution to the following variational problem. Consider the set Ψ of all piecewise differentiable functions $\phi(q)$ defined on an interval $[q_0, q_1]$, $0 \le q_0 < q_1 \le 1$, with values in the unit interval $[0,1]$, and satisfying

$$\phi'(q) + k \cdot \alpha(q, \phi(q)) \ge 0. \tag{4}$$

Let

$$T = \sup_{\phi \in \Psi} \ L(\phi) \ , \tag{5}$$

where

$$L(\phi) = 2D \int_{q_0}^{q_1} \sqrt{p(q,\phi(q))[\phi'(q) + k \cdot \alpha(q,\phi(q))]}dq. \tag{6}$$

A curve satisfying (4) is *legitimate*. Note that (4), together with the non-negativity of the distribution function p, guarantee that the square root in (6) is defined. (We always take the positive root). For a legitimate curve ϕ, the non-negative value $L(\phi)$ is the *length* of the curve.

The constrained variational problem above has been studied previously, mostly in the context of Lorentzian geometry. In the Lorentzian geometry literature, legitimate curves are called *timelike*, and the Euler-Lagrange equation associated with the functional L is an example of a *geodesic equation*. Here are a couple of basic, well-known facts about this problem, which can be found in Sections 6 and 7 in Penrose (1972).

Fact 1: The supremum of the functional (6) over all curves satisfying (4) is always achieved; hence T is the maximum of $L(\phi)$ over all curves satisfying (4). Such a curve, which maximizes $L(\phi)$, is a *maximal curve*.

Fact 2: Maximal curves are differentiable at all points in which p is differentiable and non-degenerate, $p \neq 0$, including boundary points. At such points it is also true that the inequality (4) is strict.

The analysis carried out in Bachmat *et al.* (2005, 2006) suggests that, as the number of passengers n becomes large, $t \approx T\sqrt{n}$. More precisely, for any $\varepsilon > 0$, with probability approaching 1 as $n \longrightarrow \infty$,

$$1 - \varepsilon < \frac{t}{T\sqrt{n}} < 1 + \varepsilon. \tag{7}$$

Consequently, we will use $T = T(p, k)$ as our normalized measure for boarding time.

The analysis suggests, furthermore, that, with probability approaching 1 as $n \longrightarrow \infty$, the curve obtained by joining points along the critical path by straight line segments will lie in an ε-neighborhood of a maximal curve. The maximal curves represent therefore the limiting shapes of the critical paths.

4 Airplane boarding policies

Our aim is to explain how airline boarding policies can be compared using our model. In this section we introduce the airline boarding policies which will be analyzed in this paper, and explain how to assign to each of them a joint distribution p which will serve as an input parameter for the model.

4.1 Announcement policies

An *announcement policy* is a policy in which the passengers are divided into groups $G_1, ..., G_m$. The groups are then called to board the airplane, starting with the passengers of the first group, then the second group and so on, until finally those of the m'th group are called to board the airplane. We assume that passengers are attentive to the announced order of boarding and therefore that all passengers from group G_i board before passengers from group G_j if $i < j$. We also assume that the policy does not provide any further control information beyond the specification of the groups. Hence, the boarding order within a group is uniformly random.

An important special case of announcement policies is the family of *row policies*, in which the groups are specified by blocks of contiguous row numbers. More explicitly, we consider a partition of the unit interval $0 = r_1 < r_1 < \ldots < r_m < r_{m+1} = 1$ and a permutation σ on $1 \ldots, m$. Recall that there are n/h rows in the airplane. The group G_i consists of all passengers from row $r_{\sigma(i)}(n/h)$ to row $r_{\sigma(i)+1}(n/h)$, namely, those whose normalized row number r satisfies $r_{\sigma(i)} \leq r < r_{\sigma(i)+1}$. If σ is the permutation $m, m-1, \ldots, 1$, the announcement policy is a *back-to-front policy*, since it boards passengers from the back of the airplane first, progressively moving towards the front. Such policies are very common in practice.

An announcement policy is *uniform* if all groups of passengers are of equal size, i.e., if $r_i = \frac{i-1}{m}$ for some m. We denote by F_m the uniform back-to-front announcement policy with m groups.

More generally, we denote by $F_{m,\sigma}$ the uniform announcement policy with m groups and the permutation σ.

As an example, the policy F_1 corresponds to having no policy. The order in which passengers queue is uniformly random. Let σ be the identity permutation on two elements. Then $F_{2,\sigma} = F_{2,id}$ is the policy which first allows passengers from the front half of the airplane to board, followed by passengers from the back half. By contrast, if σ is the permutation $(2, 1)$, then $F_{2,\sigma}$ is the policy which first queues passengers from the back half of the airplane followed by passengers from the front half. This is an example of a back-to-front policy. If $\sigma = (2, 1, 3)$, then $F_{3,\sigma}$ denotes the policy which allows the middle third of the airplane to board first, followed by the front third and finally the back third.

4.2 Multiclass policies

More general airline boarding strategies can be described as tiered systems consisting of classes C_i, which are not formed of contiguous blocks of rows. The classes C_i are further subdivided into groups $G_{i,j}$. The boarding order is determined by the class and the group within the class. Such strategies have been examined in Van Landeghem and Beuselinck (2002), Van den Briel *et al.* (2003) and Ferrari and Nagel (2005). We provide two examples of classes.

- Half-row classes: These classes were introduced in Van Landeghem and Beuselinck (2002). There are two classes, consisting of passengers on the right side of the aisle and those to the left of the aisle, respectively. Note that in this 2-class system, unlike the seat type classification system presented below, the delay distribution remains unaffected.

- Seat type classes: Passengers may be classified according to seat type, instead of row type. We can group all window seat passengers in one class, middle seat passengers in a second class and aisle seat passengers in a third class. Seat type classification has a feature which makes it more advantageous than other types of classification. The delay distribution D is greatly reduced by allowing window passengers to board first, followed by middle seat passengers, since passengers do not need to unseat others who are seated already.

We will consider in this paper multiclass policies with an equal number of passengers in each class. In terms of notation, consider a policy with c classes and m groups per class. We denote by $G_{i,j}$ the group consisting of passengers from the j'th class which are from rows $(i-1)/m \leq r \leq i/m$. We also use the single index notation $G_{i+(j-1)m} = G_{i,j}$. Let σ be a permutation of the elements $1, \ldots, cm$. We denote by $F_{c,m,\sigma}$ the policy which calls passenger groups in the order given by σ.

As an example, if the passengers are divided into two classes according to half rows and $\sigma = (2, 1, 4, 3)$, then the policy $F_{2,2,\sigma}$ boards first passengers to left of the aisle in the back half, followed by passengers which are left of the aisle but in the front half, then passengers on the right of the aisle in the back half and finally passengers on the right of the aisle in the front half. On the other hand, the policy with $\sigma = (2, 3, 4, 1)$ first boards passengers which are on the left in the back half, then on the right in the front half, then on the left in the back half and finally on the right in the front half.

Given a permutation σ on m elements and c classes, there is a simple way of constructing a class policy. We simply perform the $F_{m,\sigma}$ policy on the passengers of class 1, then on those of class 2 and so on. We call such a policy a *cyclic* class policy and denote it by $F_{c,m,\sigma}$.

As an example, the half-row policy with $\sigma = (2, 1, 4, 3)$ is the same as the policy $F_{2,2,\tau}$, where $\tau = (2, 1)$.

4.3 Modeling policies

In order to compare boarding policies, we need to show how to assign a distribution $p(q, r)$ to the policy. We first consider row policies. Let $q_i = r_{\sigma(i)+1} - r_{\sigma(i)}$ be the size of the block of rows of group G_i. Let $s_i = \sum_{j=1}^{i-1} q_i$ be the sum of block sizes of all passengers in groups $G_1, \ldots, G_{i-1}$. We will assume that the airplane is full. In that case the number of passengers in a group is proportional to the number of rows in the block defining the group. From our assumptions that passengers obey the policy rules it follows that pairs (q, r) which represent passengers from group G_i will form a square S_i in the unit square given by $s_i \leq q < s_{i+1} = s_i + q_i$ and $r_{\sigma(i)} \leq r < r_{\sigma(i)+1} = r_i + q_i$. The squares S_i for the permutations $\sigma = (4, 2, 3, 1)$ and $\sigma = (4, 3, 2, 1)$ are depicted in figures 4.3 and 4.3. By our assumption that boarding is otherwise uniformly random, $p(q, r)$ restricted to S_i is constant. The assumption that the airplane is full means that each set of rows has the same number of passengers and

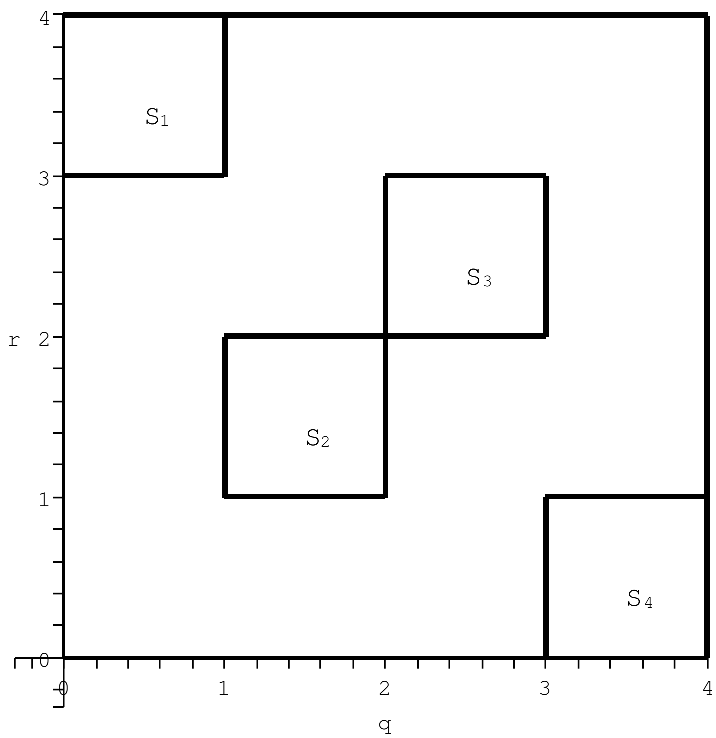

Figure 1: The squares S_i for the permutation $\sigma = (4, 2, 3, 1)$.

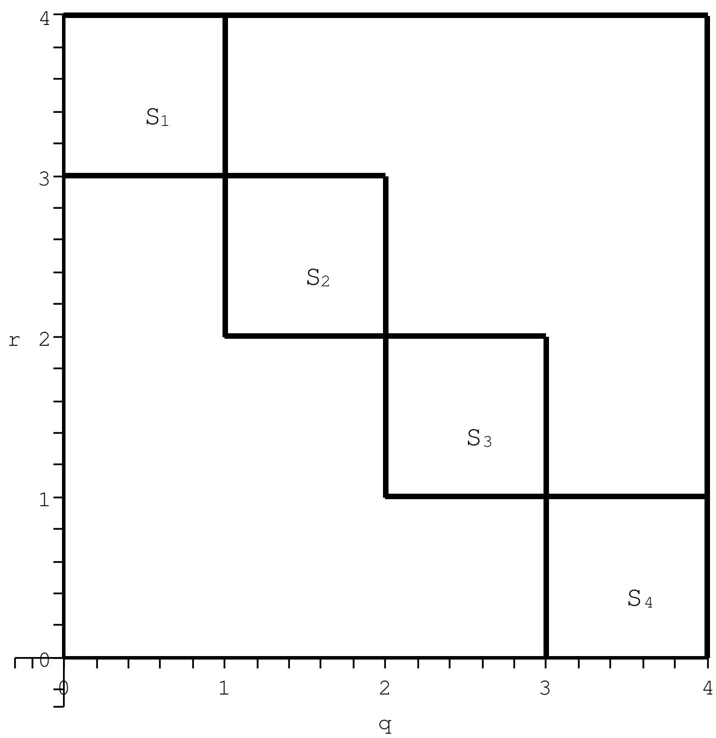

Figure 2: The squares S_i for the permutation $\sigma = (4, 3, 2, 1)$.

consequently $\int_0^1 p(q,r)dq = 1$ for all $0 \le r \le 1$. Since p vanishes outside the squares S_i, we get that $p(q,r)$ restricted to S_i equals $1/q_i$. A point (q,r) with $s_i \le q < s_{i+1} = s_i + q_i$ is below the square S_i if $r < r_{\sigma(i)}$. For such points $p(q,r) = 0$. Consequently, any segment ψ of a curve ϕ which passes entirely beneath the squares s_i satisfies

$$L(\psi) = 0 \tag{8}$$

and does not contribute to $L(\phi)$. Moreover,

$$\alpha(q,r) = \int_r^1 p(q,z)dz = \int_{r_{\sigma(i)}}^{r_{\sigma(i)+1}} 1/q_i dz = 1.$$

Plugging into (4) we see that the condition on ψ translates into

$$\psi' \ge -k. \tag{9}$$

Consider now the case of a point (q,r) with $r \ge r_{\sigma(i)+1}$ which lies above s_i. For such points again $p = 0$ and $\alpha(q,r) = 0$. For a curve segment ψ, consisting of such points, we conclude that

$$L(\psi) = 0 \tag{10}$$

and that (4) translates into

$$\psi' \ge 0. \tag{11}$$

To simplify the analysis we choose the time unit normalization $D = 1/2$, which eliminates the constant scaling prefactor $2D$ from the formula for $L(\phi)$. This procedure assumes that the boarding policies which are compared do not affect the delay, say in comparison with random boarding. This is indeed the case for row policies since the order in which passengers from the same row arrive at their seats is uniformly random.

Extending the method above to model half-row class policies is straightforward. Let $1 \le g \le cm$ be a group index and let $G_g = G_{i,j}$, which means that the g'th group to board the airplane consists of passengers from the i'th group in the j'th class, or equivalently $\sigma(g) = i + (j-1)m$. Passengers from the g'th group are represented by points in the rectangle R_g consisting of points with $(g-1)/cm \le q \le g/cm$ and $(i-1)/m \le r \le i/m$. The density p vanishes outside the union of the rectangles R_g. Also, since passengers within the g'th group board in random order, the density p is constant on R_g. Since the number of passengers in each group is the same and since all rectangles have the same size, the density p has the same values for all g. Since the density must have unit integral over the unit square and the area of the union of the rectangles R_g is $1/m$, we conclude that $p = 1/m$ on R_g for all g.

As before, points may be classified as being under, on or above R_g, and a simple computation shows that (8),(9),(10),(11) still hold.

It is important to observe that in a half-row class scheme the delay time experienced by passengers is the same as that in row policies. The reason is that the delay of a passenger depends in general only on his/her half row, since passengers in the other half do not have to get up in order for the newly arrived passenger to take his/her seat. Also the order in which passengers arrive at the half row is uniformly random as in the case of row policies. We conclude that the time unit normalization that sets $D = 1/2$ for row policies will also set $D = 1/2$ for half-row policies.

As for seat type classes, the rectangles R_i and the density p are described in the same manner, but with $c = 3$ rather than $c = 2$ for half rows. However, the delay D cannot be considered a constant anymore since passengers in different groups experience different delay values. Depending on the permutation σ, it may happen that passengers in a given block of rows are boarded using the efficient order of window first, then middle and then aisle seat. It may also happen that in other groups of rows the order is different, say, aisle, then middle and then window, which is very inefficient. In order to model this situation, we let the delay D depend on the group index. To each group G_g, consisting of the i'th group in the j'th class, we assign a delay D_g which depends on which passengers from the i'th group of rows boarded before the passengers from group G_g. As an example, the group G_g may consist of the middle passengers in the back third of the airplane. The delay of these passengers will depend on whether aisle passengers from the back third have boarded already or not. Delay times for all the different scenarios must be observed using field experiments and assigned accordingly to the different groups in the model. We then consider the delay D as a function of (q, r), setting $D(q, r) = D_g$ if $(q, r) \in R_g$ and 0 otherwise. The functional (6) should then be replaced by

$$L(\phi) = 2\int_{q_0}^{q_1} D(q, \phi(q))\sqrt{p(q, \phi(q))[\phi'(q) + k \cdot \alpha(q, \phi(q))]}dq. \tag{12}$$

5 Computations

In this section we compute asymptotic boarding times for various policies, introduced in the preceding section.

5.1 Computing the boarding time for the policy F_1

We begin our computations with the boarding policy F_1 in which passengers randomly join the boarding queue. We will recall the basic facts about this case from Bachmat *et al.* (2005, 2006) and state a few more which will serve us later.

When there is no boarding policy, the joint row/queue position distribution is uniform, namely, $p(q, r) = 1$. Since p is uniform, we have $\alpha = 1 - r$. We parameterize curves ϕ via the q-coordinate and write the curve accordingly in the form $r = r(q)$. The functional $L(\phi)$ takes the form

$$L(r) = 2\int_0^1 \sqrt{r' + k(1 - r)}dq. \tag{13}$$

Since the functional $L(r)$ does not depend explicitly on q, the Euler-Lagrange equation associated with the functional degenerates to the Beltrami equation

$$r' \frac{dL}{dr'} - L = const, \tag{14}$$

or, explicitly

$$\frac{r'}{2\sqrt{r' + k(1-r)}} - \sqrt{r' + k(1-r)} = const. \tag{15}$$

The general solution of the equation is

$$r = c_1 e^{2kq} + c_2 e^{kq} + 1. \tag{16}$$

Given such a solution in the range $\alpha \leq q \leq \beta$, the value of the length of r is

$$L(r) = (e^{k\beta} - e^{k\alpha})\sqrt{\frac{c_1}{k}}. \tag{17}$$

It is easy to see that there is a maximal curve $r(q)$ in the unit square which maximizes the functional L and satisfies $r(0) = 0, r(1) = 1$. Indeed, for any point (q, r) in the unit square the line segments connecting $(0, 0)$ to (q, r) and $q, r)$ to $(1, 1)$ are legitimate (satisfy (4)). Given a maximal curve beginning at $(q, r) \neq (0, 0)$, we can concatenate it to the line segment joining the two points to obtain a curve at least as long. The same argument applies to the point $(1, 1)$ as an endpoint. Placing these boundary conditions, we obtain the solution

$$r = \frac{e^{2kq}}{e^k - 1} - \frac{e^k}{e^k - 1} e^{kq} + 1. \tag{18}$$

The solution remains within the unit square for $0 \leq q \leq 1$ when $k \leq \ln 2$. Applying the functional to the solution, we obtain

$$T(F_1, k) = \sqrt{\frac{e^k - 1}{k}} \tag{19}$$

for $k \leq \ln 2$.

When $k > \ln 2$, the solution (18) is not contained in the unit square anymore and we have to consider the boundary of the square. We claim that the curve, which maximizes L subject to the boundary conditions $r(0) = 0,\ r(q_0) = 0,\ 0 < q_0 \leq 1$, is given by the q-axis segment $r(q) = 0$ between the endpoints. To see this, we can compare the functional L to a simpler functional $\tilde{L}(r) = \int_0^1 \sqrt{r' + k} dq$. Obviously, $\tilde{L}(r) \geq L(r)$ for any curve $r(q)$ in the unit square. It is easy to verify that the Euler-Lagrange equation associated with $\tilde{L}$ is $r'' = 0$, and so any line segment is a solution. The claim follows by noting that, for $r(q) = 0$, the values L and $\tilde{L}$ coincide.

By fact 2, a maximal curve has to be differentiable. We conclude that the maximal curve $r(q)$ consists of a segment r_1 of the form $[0, q_0]$ on the q-axis, followed by a curve r_2 which solves the Euler-Lagrange equation between the points $(q_0, 0)$ and $(1, 1)$. By the differentiability of

r we need $r_2'(q_0) = 0$. We solve the Euler-Lagrange equation with boundary conditions $r_2(q_0) = 0$ and $r(1) = 1$, requiring in addition that $r_2'(q_0) = 0$, which leads to $q_0 = \frac{k - \ln 2}{k}$. Plugging the resulting curve r into the functional leads to

$$T(F_1, k) = \sqrt{k} + \frac{1 - \ln 2}{\sqrt{k}}, \qquad k > \ln 2. \tag{20}$$

The maximal curves are plotted for various values of k in Figure 5.1.

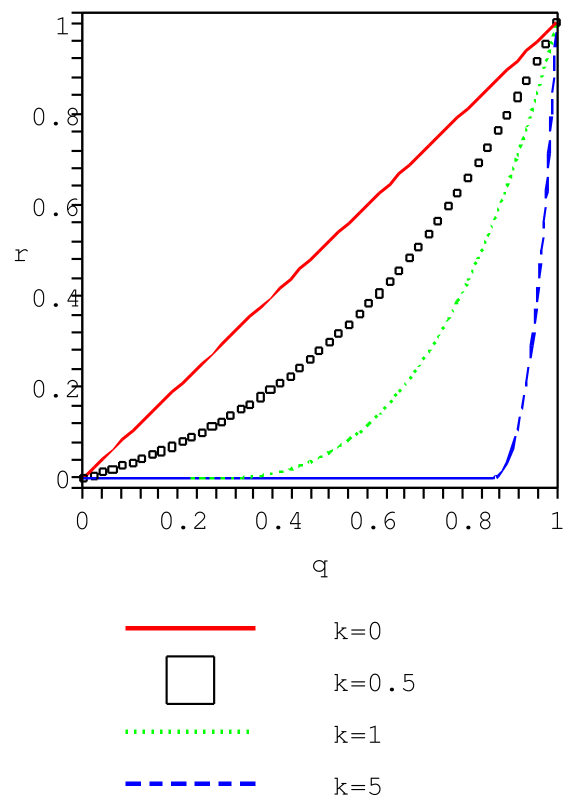

Figure 3: Maximal length curves for several values of k.

5.2 Computing boarding time for the F_2 policy

We present a general formula for $T(F_2, k)$, the normalized boarding time for the F_2 policy with parameter k, when $k \geq 1$. It is easier to study the F_2 policy if we inflate the unit square to the square $[0, 2] \times [0, 2]$ via the map $(q, r) \longrightarrow (2q, 2r)$. Let U be a maximal curve. Let U_1 be the segment of U which lies in S_1, given by $0 \leq q \leq 1$ and $1 \leq r \leq 2$. Similarly, let U_2 be the segment of U which lies in S_2, given by $1 \leq q \leq 2$ and $0 \leq r \leq 1$. We may assume that a maximal curve which has a non-empty part lying in S_1 must begin at the point $u = (0, 1)$, since the line segment connecting u with any other point v in S_1 is legitimate. By a similar argument for S_2, one of the points of the form $(1, \delta)$ with $0 \leq \delta \leq 1$ can be taken as the initial point for U_2. We also note that, when $k > 1$, the curve U_1 may be taken to be non-empty since the line segment joining u to any point of the form $(1, \delta)$ is legitimate.

The endpoint of U_1 lies on the boundary of S_1, either on the bottom edge, $r = 1$, or on the right edge, $q = 1$. By (9), curves which satisfy (4) and start at a point with $q = 1$ and $r \geq 1$ are non-decreasing and therefore never reach S_2. Consequently, the endpoint of U_1 must lie on the bottom edge of S_1. Let $(q_1, 1)$ be a point on the bottom edge. As explained in Section 5.1, the maximal curve between u and $(q_1, 1)$ is the curve $(q, 1)$, $0 \leq q \leq q_1$, and hence U_1 must have this form. The length of the curve is

$$q_1\sqrt{k/2}. \tag{21}$$

U_2, which begins at some point $(1, \delta)$, must end at $(2, 1)$ and must be maximal among all such curves. Let $U_2(\delta)$ denote the maximal curve with these endpoints. Depending on δ, it may either be a solution to the Euler-Lagrange equation, lying in the interior of S_2, or it may contain a segment on the lower edge of S_2. In either case, it must be differentiable, and hence the critical value of δ which separates the two cases is the value for which the solution to the Euler-Lagrange equation passing through $(1, \delta)$ and $(2, 1)$ is tangent to the bottom edge of S_2. The solution passing through $(1, \delta)$ and tangent to the bottom is given by the equation

$$r(q) = (1 - \sqrt{\delta})^2(e^{k(q-1)} - \frac{1}{1 - \sqrt{\delta}})^2, \tag{22}$$

and passes through the point $(2, 1)$ when $\delta = (1 - 2e^{-k})^2$. If $\delta \geq (1 - 2e^{-k})^2$, then U_2 will have the form $\frac{1-\delta}{e^k-1}e^{2k(q-1)} - \frac{e^k(1-\delta)}{e^k-1}e^{k(q-1)} + 1$. According to (16), we have

$$L(U_2(\delta)) = \sqrt{\frac{1}{2k}}\sqrt{(e^k - 1)(1 - \delta)} \tag{23}$$

in this case. If, on the other hand, $\delta < (1 - 2e^{-k})^2$, then $U_2(\delta)$ consists of three segments:

(a) the solution of (22) up to the tangent point $(1 + (\ln(\frac{1}{1-\sqrt{\delta}}))/k, 0)$,

(b) the line segment joining $(1 + (\ln(\frac{1}{1-\sqrt{\delta}}))/k, 0)$ and $(2 - \ln 2/k, 0)$, and

(c) the solution from $(2 - \ln 2/k, 0)$ to $(2, 1)$, which is tangent to the bottom edge at $(2 - \ln 2/k, 0)$ according to the computations in Section 5.1.

The length of the three segments combined is

$$L(U_2(\delta)) = \sqrt{\frac{1}{2k}}(\sqrt{\delta} + \ln(1 - \sqrt{\delta}) + k + 1 - \ln 2). \tag{24}$$

For any legitimate curve ψ in the region below S_1 we have, by (9) and (8), $L(\psi) = 0$ and $\psi' \geq -k$. The segment of U between the endpoint of U_1, which is $(q_1, 1)$, and the initial point of U_2, which is $(1, \delta)$, is legitimate, and therefore the inequality

$$1 - \delta \leq (1 - q_1)k, \tag{25}$$

must hold. It is easy to verify that, as a function of δ, the length of $U_2(\delta)$ is decreasing. Therefore, for the maximal curve U we have equality in (25). Consequently, we define $U_1(\delta)$ to be the curve given by $r = 1$ in the range $0 \leq q \leq \frac{k-1-\delta}{k}$. The length of $U_1(\delta)$ is by (21)

$$L(U_1(\delta)) = \sqrt{\frac{1}{2k}}(k - 1 + \delta). \tag{26}$$

Let $U(\delta)$ be the concatenation of $U_1(\delta)$ and $U_2(\delta)$. By the analysis presented above we know that the maximal curve has the form $U(\delta)$ for some $0 \leq \delta \leq 1$. Let $A(\delta)$ denote the sum of the expressions (24) and (26). Differentiating $A(\delta)$, we find that $A(\delta)$ increases towards $\delta = 1/4$ and is maximized when $\delta = 1/4$. The length of $U(\delta)$ coincides with $A(\delta)$ in the range $0 \leq \delta \leq (1 - 2e^{-k})^2$. For $k \geq 2\ln 2$, the value $\delta = 1/4$ is in that range and $A(1/4)$ is the maximal value, while for $k \leq 2\ln 2$ the maximal value in the range is provided by $A((1 - 2e^{-k})^2)$. Let $B(\delta)$ denote the sum of the expressions on the right-hand sides (23) and (26). $B(\delta)$ coincides with $U(\delta)$ whenever $\delta \geq (1 - 2e^{-k})^2$. Differentiating $B(\delta)$, we find that the maximal value is obtained when $\delta = 1 - \frac{e^k - 1}{4}$. This value is in the range where $U(\delta)$ coincides with $B(\delta)$ whenever $1 - \frac{e^k-1}{4} \geq (1 - 2e^{-k})^2$. Writing $x = e^k$ and multiplying both sides by x^3, which is always positive, we obtain the condition $x^3 - 16x^2 + 16x - 1 \leq 0$. Since $x^3 - 16x^2 + 16x - 1 = (x - 1)(x - 4)^2$, the condition holds for $1 \leq x \leq 4$, which translates to the condition $k \leq 2\ln 2$.

When $k \geq 2\ln 2$, the maximum of $B(\delta)$ in the range $\delta \geq (1 - 2e^{-k})^2$ is attained at $\delta = (1 - 2e^{-k})^2$. Since $A((1 - 2e^{-k})^2) = B((1 - 2e^{-k})^2)$, we conclude that for $k \geq 2\ln 2$ the maximal value of $U(\delta)$ is attained at $\delta = 1/4$. Similarly, for $k \leq 2\ln 2$ the maximal value of $U(\delta)$ is attained at $\delta = 1 - \frac{e^k-1}{4}$. Plugging these values into $A(\delta)$ and $B(\delta)$, respectively, we see that the length of the maximal curve U for $k \geq 2\ln 2$ is

$$T(F_2, k) = \sqrt{2k} + \frac{3/4 - 2\ln 2}{\sqrt{2k}} \tag{27}$$

and for $1 \leq k \leq 2\ln 2$ is

$$T(F_2, k) = \sqrt{\frac{1}{2k}}(k + \frac{e^k - 1}{4}). \tag{28}$$

The maximal curve for $k = 4$ is displayed in Figure 5.2.

5.3 Computations for F_m

Consider the case $m > 2$. As for F_2, it is more convenient to consider the expanded square $[0, m] \times [0, m]$. We have the squares $S_i = [i - 1, i] \times [m - i, m - i + 1]$, $i = 1, \ldots, m$, which lie along the anti-diagonal. The part of the maximal length curve U which contributes to the length is composed as before of segments U_i contained in S_i. In addition, there are legitimate segments, joining the endpoint of U_i to the initial point of U_{i+1}, which lie below the S_i and do not contribute to the length. The arguments presented for the case of F_2 also show that $(0, m - 1)$ is the initial point of U (and U_1) and that the initial point of U_i is on the left edge

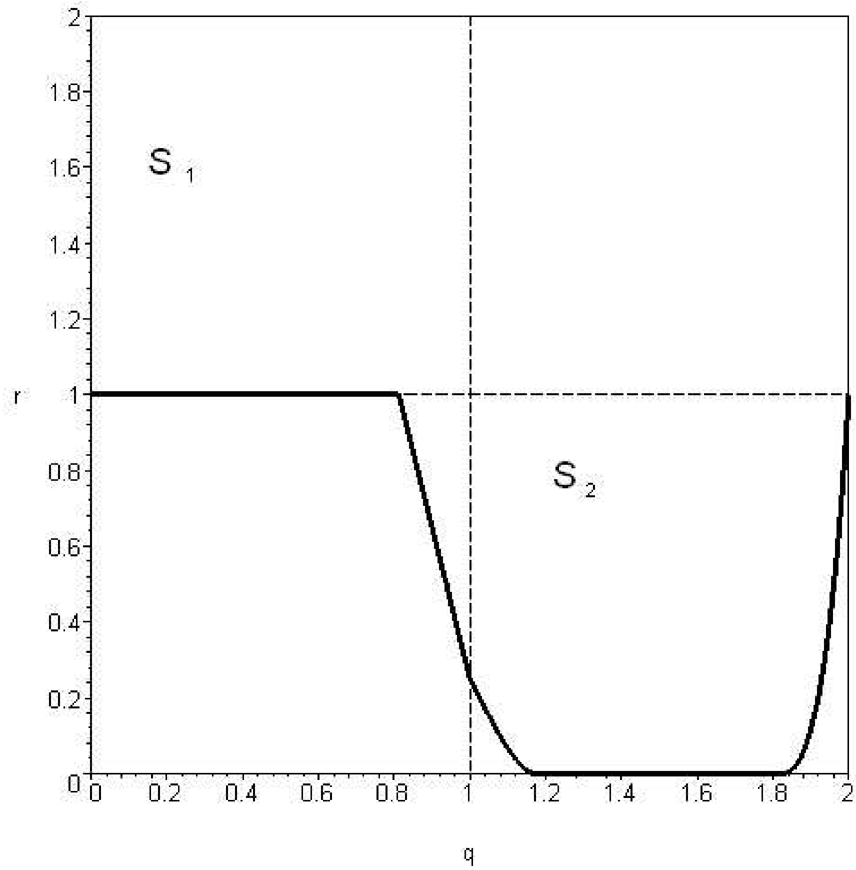

Figure 4: Maximal curve for F_2 with $k = 4$.

of S_i for all i. In addition, for $i = 1, \ldots, m-1$, the endpoint of U_i lies on the bottom edge of S_i. Thus, the initial point of U_i has the form $(i, m-i+\delta_i)$ and the endpoints have the form $(i+\beta_i, m-i)$ for some δ_i and β_i. The maximality of U, together with the requirement that U be legitimate, implies the relation

$$\beta_i = 1 - \frac{1-\delta_{i+1}}{k}. \tag{29}$$

Consider U_1 and U_2. The initial point and the endpoint of U_1 lie on the bottom edge of S_1, and hence by maximality all of U_1 lies on the bottom edge. Let $(1, m-2+\delta)$ be the initial point of U_2 and $(1+\beta, m-2)$ the endpoint. Given the endpoint of U_2, we can optimize the location of the initial point of U_2. Depending on δ_2, U_2 may either be a solution to the Euler-Lagrange equation which lies in the interior of S_2 or be composed of a solution which is tangent to the bottom edge, followed by a segment along the bottom edge. In the first case the length of the segment is

$$L(U_2) = \frac{\sqrt{e^{kb}-1}}{\sqrt{k}} \cdot \frac{\sqrt{(1-\delta)e^{kb}-1}}{\sqrt{e^{kb}-1}}. \tag{30}$$

We may add the contribution of U_1 and differentiate. The sum of contributions is maximized when

$$\delta = -(\frac{e^{kb}}{4} + \frac{4}{e^{kb}}) + \frac{5}{4}. \tag{31}$$

Writing $x = \frac{e^{kb}}{4}$, we see that $\delta = -(x+1/x)+5/4$, and since $x+1/x \geq 2$ for all $x > 0$ we see that the optimal δ is negative. Hence the U_2 component of the maximal curve is composed of a solution tangent to the bottom of S_2, followed (possibly) by a segment along the bottom. This case was already analyzed in the F_2 case. The optimal value for δ is $1/4$ if it is in the range

$$\ln(\frac{1}{1-\sqrt{\delta}})/k \leq b. \tag{32}$$

Let $k \geq 3/4 + \ln 2$. Consider the curve U with $\delta_i = 1/4$ for $i = 1, \ldots, m-1$. By (29), this corresponds to $\beta_i = 1 - \frac{3}{4k}$. We note that (32) is satisfied for $b = \beta_i$ when $k \geq 3/4 + \ln 2$. We claim that U is maximal. Assume to the contrary that U' with parameters β'_i is maximal. Let j be the first index for which $\beta'_j \neq 1 - \frac{3}{4k}$. We claim that $\beta'_j < 1 - \frac{3}{4k}$, as otherwise, by the calculations for U_1 and U_2, the path U'_{i+1} from $(j+1, m-(j+1)+\delta'_{j+1})$ to $(j+1+\beta'_{j+1}, m-(j+1))$ is not optimal, since $\delta'_{j+1} \geq 1/4$. For the same reason we must have $\beta'_i < 1 - \frac{3}{4k}$ for all $i > j$. However, by the computation for F_2 we know that for $k \geq 2\ln 2$ the contribution of the union of U_{m-1} and U_m is not optimal. Since $3/4 > \ln 2$, we are done.

The length of the curve U is

$$T(F_m, k) = \sqrt{mk} - \frac{m-2}{\sqrt{mk}}(\ln 2 + 1/4) - \frac{2\ln 2 - 3/4}{\sqrt{mk}}. \tag{33}$$

The maximal curve for F_3 with $k = 4$ is depicted in Figure 5.3. The maximal curve for F_m has the same structure with the segments $U_2, \ldots, U_{m-1}$ repeating.

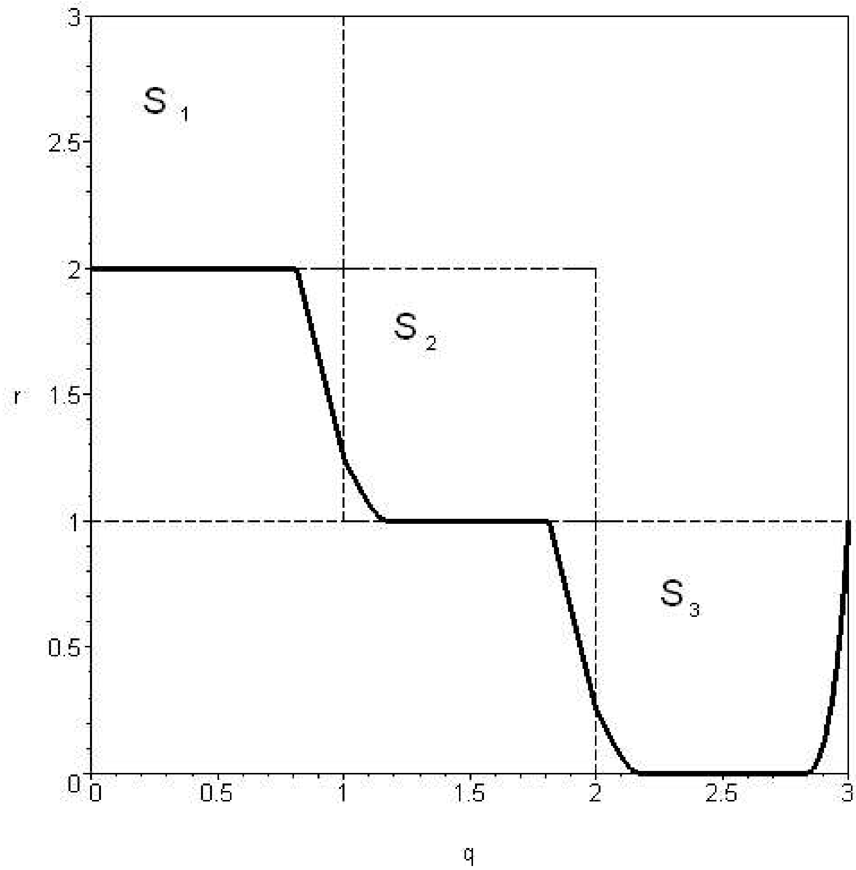

Figure 5: Maximal curve for F_3 with $k = 4$.

5.4 Computing $F_{m,\sigma}$

We compute $F_{m,\sigma}$ under some assumptions on σ and k. Using the representation of passengers in the expanded square $[0,m]\times[0,m]$, elements in the i'th group will be represented by points in the square $S_{i,\sigma}$, consisting of points of the form $i-1\leq q\leq i$ and $\sigma(i)-1\leq r\leq\sigma(i)$. Let U be a maximal curve, composed of curves U_i in $S_{i,\sigma}$. The sequence $\sigma(1),\sigma(1),\ldots,\sigma(m)$ decomposes into blocks B_j, $1\leq j\leq s(\sigma)$ of size b_j of decreasing sequences, that is, $\sigma(1)>$ ... $>\sigma(b_1)<\sigma(b_1+1)>\ldots>\sigma(b_1+b_2)<\sigma(b_1+b_2+1)\ldots$. Let $c_j=\sum_{i=1}^{j}b_j$ be the sequence of partial sums of the sequence b_j. Define the *excess descent* of block B_j by $e_j=\sigma(c_j)-\sigma(c_{j-1})-b_j+1$ if $b_j>1$ and $e_j=0$ if $b_j=1$. Let $e=\sum_j e_j$. Let $\lambda(\sigma)$ be the maximum over all decreasing pairs of consecutive elements, $\sigma(l)<\sigma(l+1)$, of $\sigma(l+1)-\sigma(l)$. If

$$k\geq 3/4+\ln 2+\lambda(\sigma)-1, \tag{34}$$

then

$$T(F_{m,\sigma},k)=\sum_{j=1}^{s(\sigma)}T(F_{b_j},k)\frac{\sqrt{b_j}}{\sqrt{m}}-e\frac{1}{\sqrt{km}}. \tag{35}$$

To see this, consider a maximal causal curve U in the union of $S_{i,\sigma}$, composed of curves U_i in the respective squares, and legitimate linear segments joining them. As in the previous computations, it is easy to verify that the endpoint of U_i is on the bottom edge of $S_{i,\sigma}$, say at $(i-1+\beta_{i,\sigma},\sigma(i)-1)$, and the initial point of U_i is on the left border of $S_{i,\sigma}$, say at $(i-1,\sigma(i)-1+\delta_{i,\sigma})$. The endpoint of U_i must be connected to the initial point of U_{i+1} by a legitimate linear segment. If $\sigma(i)<\sigma(i+1)$, then the line segment between any point of $S_{i,\sigma}$ and any point of $S_{i+1,\sigma}$, is legitimate and therefore we can make independent computations in the blocks B_j as long as the maximal curves in B_j contain non-empty curves U_i for all i. When $\sigma(i)>\sigma(i+1)$, (4) implies that

$$\sigma(i)-\sigma(i+1)-\delta_{i+1}\leq k(1-\beta_i). \tag{36}$$

The process of identifying a maximal curve in B_j is identical to the process in the case of F_{b_j}, subject to the replacement of (29) by (36) and a change in density from $\frac{1}{b_j}$ to $\frac{1}{m}$. As in the case of F_{b_j}, the curve with $\delta_i=1/4$, for $c_{j-1}+2\leq i\leq c_j$, is optimal when it can be constructed subject to (36), which holds when $k\geq\sigma(i)-\sigma(i+1)-1+3/4+\ln 2$ for all i. The change in density yields a comparison with $T(F_{b_j},k)\frac{\sqrt{b_j}}{\sqrt{m}}$. A comparison of (29) and (36) shows that the portion of the curve U_i on the bottom edge of $S_{i,\sigma}$ is shorter (in the standard Euclidean sense) by $\frac{\sigma(i+1)-\sigma(i)-1}{k}$ than the corresponding curve in the computation of $T(F_{b_j},k)$. Summing over $i\in B_j$ and taking the density $\frac{1}{m}$ into account, we obtain the difference term $-e\frac{1}{\sqrt{km}}$, thus establishing (35).

5.5 Computing $F_{2,m,\sigma}$

We compute boarding time for cyclic half-row policies $F_{2,m,\sigma}$ for certain permutations σ. Consider the restriction of the boarding process to the first half of the queue. We wish to compare this part of the boarding process to the boarding process associated with the policy $F_{m,\sigma}$. The number of passengers is $n/2$ instead of n. Also, h should be replaced by $h/2$, since only half the number of passengers in each row (those to the left of the aisle) are boarding. The order in which row blocks board is by definition the same as that of $F_{m,\sigma}$. This leads to the replacement of k by $k/2$. We conclude that the boarding time for the first half of the queue T_1 is given by

$$T_1 = T(F_{m,\sigma}, k/2)\sqrt{1/2}\,. \tag{37}$$

By symmetry, the same formula would hold for the boarding time if we considered only the second half of the queue. Assume that $\sigma(1) < \sigma(m)$. This means that each passenger from the last group of the first half of the queue blocks each passenger from the first group in the second half of the queue. Assume furthermore that k is large enough so that the maximal curve U for half the queue contains non-empty segments in the first and last groups, namely U_1 and U_m are non-empty when considering the policy $F_{m,\sigma}$ with congestion parameter $k/2$. We conclude that in such cases the maximal paths for the first and second half of the queue can be concatenated to produce a maximal curve of twice the length. Therefore, under such circumstances

$$T(F_{2,m,\sigma}, k) = 2T_1 = \sqrt{2}T(F_{m,\sigma}, k/2). \tag{38}$$

6 Comparison with previous work

The model and the computations presented above must be validated against detailed simulations, since they make many simplifying assumptions. In this section we compare our computations with simulation results reported in previous works on airplane boarding, in particular with those of Van Landeghem and Beuselinck (2002), which are the most detailed. Van Landeghem and Beuselinck (2002) have carried out computer based simulations of 47 different boarding policies.

The design and input of the simulations were based on observations of passenger boarding and interviews with personnel at Brussels airport. Each experiment was performed 5 times, and the average and standard deviation are recorded on page 302 of loc. sit. and plotted on page 303. A detailed description of the boarding procedures which were simulated is given on pages 299–300. We shall refer to the simulations as the V-B simulations.

The simulations take into account some observed delays which we have not modeled. We briefly summarize the differences between the V-B simulations and our model.

- Passengers in the V-B simulations travel along the aisle at finite speeds according to some distribution. In our boarding model, we implicitly assume that passengers travel at infinite speeds when unobstructed. This leads to the definition of delay as the time spent while the passenger is blocking the aisle at his/her row location.

- The V-B simulations attach to each passenger between one and three carry-on items, to be stored in the overhead bin compartments. The simulation keeps track of the available bin space. Towards the end of the boarding process, passengers sometimes need to search a large number of bins before they find space for their luggage, causing further delay in aisle clearing time. Also, as noted earlier, late arriving passengers have to unseat other passengers in order to get to their assigned seat. Consequently, the delay becomes larger towards the end of the V-B simulation. As a result of these considerations, the delay in the V-B simulations depends on the (q, r) coordinates of the passenger and is also given by a distribution rather than a constant value as assumed in the model.

- The simulations also take into account the situation in which an occasional passenger sits in the wrong place. When the passenger whose assigned seat is taken arrives, the passenger who occupies the wrong seat gets up and moves to the correct place. This phenomenon is not considered at all in the model.

- The V-B simulations model an Airbus A320 with 23 rows and 132 seats, each row (apart from first class) carrying six passengers, three on either side of the single aisle. Our computations are asymptotic and assume that the number of passengers is very large. As shown in Bachmat *et al.* (2006) the computations are not accurate at all for 100-200 passengers per aisle, a reasonable range for the number of passengers per aisle in a modern commercial airplane. As an example, we refer to Figures 6 and 7. According to the model, the critical path for the airplane boarding process should follow closely the shape of the maximal curve. In the two figures the maximal curve is plotted in red, while the other curves display the critical path for runs of the airplane boarding project with 100, 200 and 10^6 passengers. in Figure 6 we see the results for runs with $k = 0.5$, while in Figure 7 we see the results for $k = 5$. As can be seen, when there are 10^6 passengers the critical paths do follow closely the maximal curve, while for 100 and 200 passengers there are very large differences. These differences also display themselves in the differences between the actual length of the critical path and the asymptotic calculations. The analysis of these differences in Bachmat *et al.* (2006) shows that the asymptotic expressions of the form $T(F, k)\sqrt{n}$ consistently overestimate the boarding time. The overestimation increases with k, and for reasonable values of k such as $k = 4$ is on the order of 50% .

  However, we are interested in comparing different boarding policies. When comparing two different policies we consider the ratio of boarding times. Consistent overestimation will have little effect on such ratios. Thus, while the model based asymptotic estimates

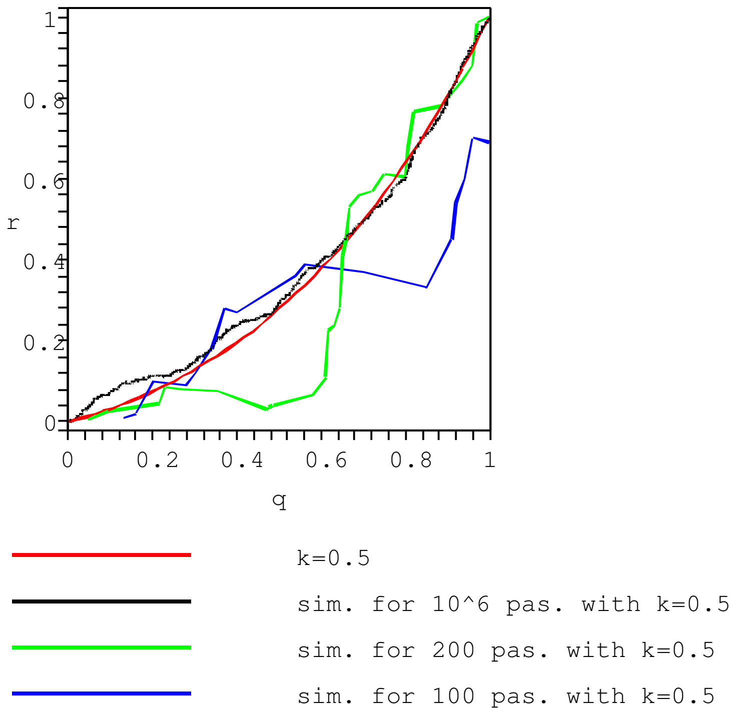

Figure 6: Comparison of Maximal curve for k=0.5 with simulation based maximal chains.

are not reliable indicators of absolute boarding times, they may still be considered for policy comparisons.

Having stated the difference between the V-B simulations and our model based computations, we proceed to the comparison of their respective results. In the V-B simulations, the 132 passengers are divided into groups called in succession to join the queue, while the order within the group is random. In 18 of the experiments, the number of groups is very large and the boarding process becomes nearly deterministic, corresponding to various combinatorial instances of the boarding process, as discussed in Section 2. The fastest boarding methods according to the simulations belong to these tightly controlled boarding methods. The best method is the one calling window passengers from one side of the aisle first in descending order, followed by window passengers from the other side in descending order, and similarly for middle and then aisle passengers. The problem of families or other small parties of passengers traveling together being separated by such policies can be solved by allowing such parties to board together according to the minimal group number among participants. Van Landeghem and Beuselinck found this policy to be about 2.5 times faster than random boarding. This deterministic policy is also the fastest according to our tasks with precedences model, allowing passengers to board in 6 rounds, far better than the random policy. In fact, the relative ranking among these 18 policies nearly coincides with the total boarding time according to

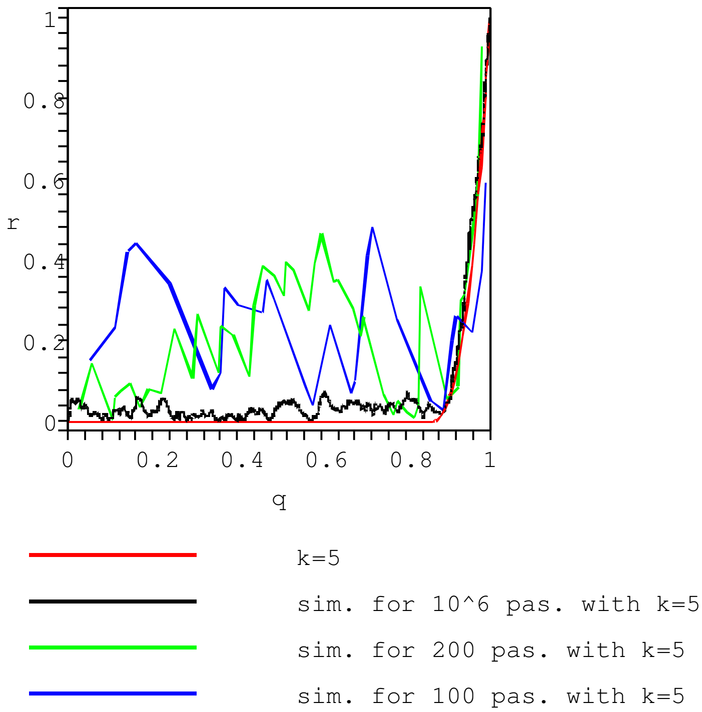

Figure 7: Comparison of Maximal curve for k=5.0 with simulation based maximal chains.

our corresponding deterministic tasks with precedence models. It is unlikely, though, that an airline can exert such detailed control on the order in which passengers board, and we do not know of any example in which such strict methods are practiced. Therefore, we will not analyze these policies in great detail.

The remaining 29 simulations concern row assignment policies blocks or half-row class policies. We can compare the results of these simulations with the calculations we have performed. We omitted one policy which involves placing first-class passengers first and, in addition, three policies which divided the passengers into 20 groups. As noted earlier, such policies are hard to enforce, and, in addition, for such a large number of groups the probabilistic analysis tends to be irrelevant. This leaves us with 25 policies. We need some estimate of the parameter k. Given that there are 6 passengers per row, and assuming an estimate of 2/3 of the distance between successive rows for the average aisle length occupied by a passenger, we obtain an estimate of $k = 4$. We performed a complete set of computations assuming this value for the 25 policies considered by Van Landeghem and Beuselinck. We also computed boarding times for the policies with the estimate $k = 3.5$, and observed that the results remain qualitatively the same. Almost all computations were performed using (27), (28), (33), (35) and (38) in a straightforward manner. In a few cases, some additional arguments were needed. An explanation of the basis of these other calculations is given in

Appendix B.

In terms of our terminology, Van Landeghem and Beuselinck simulated F_1, F_2, F_3, F_4, F_6 and F_{10}. In addition, they simulated $F_{2,2}$, $F_{2,3}$, $F_{2,4}$, $F_{2,6}$ and $F_{2,10}$, using passengers on the right side of the aisle as the first class and passengers on the left side as the second class. The remaining 15 out of 25 simulations were of type $F_{m,\sigma}$ and $F_{2,m,\sigma}$ for various permutations. A list of the permutations used in the V-B simulations is given in Appendix A.

We consider F_1 to be the basic policy which will serve as a yardstick for measuring all other policies, and normalize its boarding time to be 1 (24.7 minutes according to loc. sit.). We compute for each policy the ratio of the boarding time of the policy to that of F_1. The computations are done with the parameter setting $k = 4$. The results are presented in Figure 6 and table 1.

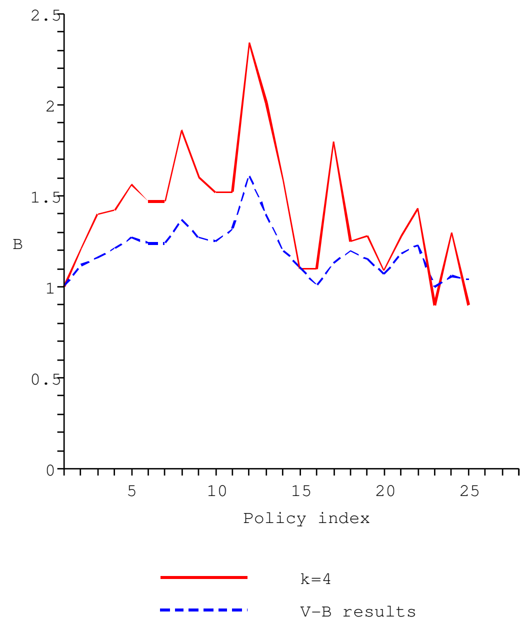

Figure 8: Boarding time (B) according to the model and the V-B simulations.

We make the following observations and conclusions.

A) The computational results using the model based approach and the results of the V-B simulations agree to a large extent. The correlation coefficient between the $k = 4$ computations and the V-B results is 0.965. There is also strong agreement with the V-B simulations on the ordering of policies according to boarding time, as can be seen in Table 1, which provides the comparison of policy rankings. This result – we believe – reinforces both approaches.

B) The main disagreement between results comes in cases where the number of classes is large. In experiment 12 with policy F_{10}, we are in a situation with $k > 2\ln 2$ and $m = 10$.

index	policy	$k = 4$	V-B results	$k = 4$ ranking	V-B ranking
1	F_1	1.00	1.00	3	1
2	F_2	1.21	1.12	8	8
3	F_3	1.40	1.16	13	11
4	F_{3,σ_0}	1.42	1.21	14	15
5	F_4	1.56	1.27	20	20
6	F_{4,σ_1}	1.47	1.24	16	17
7	F_{4,σ_2}	1.47	1.24	17	18
8	F_6	1.86	1.37	23	23
9	F_{6,σ_3}	1.60	1.27	22	21
10	F_{6,σ_4}	1.52	1.25	18	19
11	F_{6,σ_5}	1.52	1.32	19	22
12	F_{10}	2.34	1.61	25	25
13	F_{10,σ_6}	2.01	1.40	24	24
14	F_{10,σ_7}	1.58	1.20	21	14
15	$F_{2,2}$	1.10	1.11	5	7
16	$F_{2,2,\sigma_8}$	1.10	1.01	6	3
17	$F_{2,3}$	1.18	1.13	7	9
18	$F_{2,3,\sigma_9}$	1.25	1.20	9	13
19	$F_{2,4}$	1.28	1.15	10	10
20	$F_{2,4,\sigma_{10}}$	1.09	1.07	4	6
21	$F_{2,4,\sigma_{11}}$	1.28	1.18	11	12
22	$F_{2,6}$	1.43	1.23	15	16
23	$F_{2,6,\sigma_{12}}$	0.90	1.00	1	2
24	$F_{2,6,\sigma_{13}}$	1.30	1.06	12	5
25	$F_{2,6,\sigma_{14}}$	0.90	1.04	2	4

Table 1: Comparison results

Each class in this case contains merely 13 passengers. By (33), the boarding time of F_m tends to infinity with m, resulting in an overestimate. Underestimates occur when there are many small groups which do not block each other as in experiments 23 and 25.

C) According to (33) and (38), for fixed values of m and $k \geq 2(3/4 + \ln 2)$ we have $T(F_{2,m}, k) < T(F_m, k)$. In agreement with this statement, we observe that indeed the boarding times of F_2, F_3, F_4, F_6 are all greater than the corresponding boarding times of $F_{2,2}, F_{2,3}, F_{2,4}, F_{2,6}$, according to the V-B simulations.

D) In some cases the V-B simulations distinguish between pairs of policies which our models cannot distinguish. In the pairs 15 and 16, 19 and 21, and 23 and 25, the groups of passengers consist of passengers from the same blocks of rows but from different sides of the aisle. Our models of airplane boarding cannot distinguish between seat assignments differing only by a symmetry with respect to the aisle, and hence the expected boarding time is the same for the pairs. In all three pairs, the V-B results show a difference. This difference is small for 19 and 21, as well as for 23 and 25, but amounts to 10 percent for the pair 15 and 16. The most natural explanation for the difference is simple statistical fluctuations of the average over the 5 trials of each experiment. The average of the experiments turns out to be in very good agreement with the analytical computations. Another possible explanation is the fact that the V-B simulations keep track of which overhead bins are available. If the simulations also assume that passengers who sit in a given side try to place their luggage on that same side, then the symmetry is broken and the experiments can be distinguished, even though the differences should still be small.

E) Given observation (C), it is natural to suggest the use of policies with 3 classes. The division of passengers into classes may be random, or may be according to the row number modulo 3. We can then impose an F_m type policy in each class in succession. Such class policies, which may have in general c classes, can be modeled in the same manner that we have modeled half-row classes above. The same arguments which are used to establish formula (38) can be used to show that for $k \geq c(3/4 + \ln 2)$ the model based boarding time estimate is

$$T(c, m, k) = \sqrt{c}\, T(m, k/c)\,. \tag{39}$$

In particular, we may consider $F_{3,2}$, which divides passengers into 3 classes and 6 groups. By observation (A), we expect our predictions for this policy to be fairly good since m and the number of groups cm are small and $k/c > 1$. The expected boarding time is 0.95 when we assume $k = 3.5$, 0.99 when we assume $k = 4$ and 1.04 when we assume $k = 4.5$. It seems unlikely that passengers will tolerate a division into more than 6 groups, and within that range $F_{3,2}$ has the best predicted performance. We note that the expected gain over the uniform policy F_1 is rather negligible and is probably not worth the extra complication. We conclude, in agreement with the simulation results of V-B, that it is difficult to improve upon the uniform boarding policy, using policies which do not change the delay distribution D, without excessively burdening the passengers.

F) We are left with the option of using seat type classes which decrease the value of D and provide a better way of dividing passengers into 3 classes. The policy $F_{3,1}$ with seat type classification, along with some variants which resemble $F_{3,2}$, have been suggested by Van den Briel *et al.* (2003, 2005). We note that $F_{3,1}$ simply means that window seat passengers are boarded first, followed by middle seat passengers and finally aisle seat passengers. This policy is also sometimes referred to as an *outside-in* policy. Van den Briel *et al.* consider a different measure for assessing airplane boarding policies. It can be shown that asymptotically their measure correlates strongly with the number of times that passengers which are already seated have to get up to allow newly arrived passengers reach their seat. It is therefore not surprising that seat type classes and other outside-in boarding policies prove to be optimal with respect to this measure.

Van den Briel *et al.* (2005) report on the success of a change in policy at America West airlines. They converted from a (non-uniform) back-to-front policy to a policy which which they call *reverse pyramid.* Reverse pyramid closely resembles $F_{3,2}$ with seat type classification, but has only 5 groups. A 20 percent reduction in boarding time is reported. According to our analysis $F_{3,1}$, inverted pyramid and $F_{3,2}$ should improve upon F_1 by about the same amount, with a slight edge for $F_{3,2}$ over $F_{3,1}$. $F(3,1)$ and inverted pyramid have been considered in Ferrari (2005), which simulates many policies. They have been found to provide a 21 and 25 percent reduction in boarding time, respectively. This shows again that the main gain in boarding time comes from the employment of outside-in strategies rather than from experimenting with the ordering of rows.

7 Announcement policies and congestion

Having validated the model, we would like to explore some of its consequences. We show that the relative effectiveness of announcement policies changes substantially between small and large values of k. When $k = 0$, passengers from different groups in a back-to-front policy do not interact with each other during the boarding process. Consequently, the normalized boarding time is the maximum among boarding times in each individual group. The i'th group has $q_i n$ passengers and hence its normalized boarding time is $\sqrt{q_i}$ times faster than the boarding time with no policy. The normalized boarding time is thus $\max_i \sqrt{q_i}$. We conclude that the policy with worst boarding time among back-to-front policies is the uniform policy F_1 and that, among all announcement policies (not only back-to-front) with a fixed value of m, the uniform policy F_m is the best. On the other hand, the following result shows that, as k becomes large, the uniform policy F_1 beats any other announcement policy, and among all policies with a fixed value of m the uniform policy F_m is asymptotically one of the worst policies and is worst among back-to-front policies. This shows again the important observation that congestion is a critical parameter in airplane boarding. In addition, the theorem shows that as congestion increases the order in which groups are boarded becomes less important and the various sizes of the groups are the only parameters that counts. In the following theorem we denote by $T(F, k)$ the normalized boarding time of a policy F with parameter k.

Theorem 7.1 *For any announcement policy F, given by $r_1, \ldots, r_{m+1}$ and a permutation σ, with $m > 1$, there exists a value k_F such that for $k > k_F$ we have $T(F_1, k) < T(F, k)$. More precisely,*

$$\lim_{k\to\infty} \frac{T(F,k)}{\sqrt{k}} = \sum_i \sqrt{q_i},$$

and, in particular, the limit is independent of σ.

Proof: Let $q_{min} = \min_i \; q_i$. A curve satisfying (4) consists of a union of curves U_i in the rectangles R_i, which describe the boarding process among passengers in rows $r_{i+1} \le r \le r_i$, and of curves which do not contribute to (6), but which satisfy (4) and connect the endpoint of U_i to the initial point of U_{i+1}.

By the computation of Section 5.1, applied to the nq_i passengers in the i'th group, we have $L(U_i) \le \sqrt{q_i k} + \sqrt{q_i k}\frac{1-\ln 2}{k} = \sqrt{q_i k} + o(1)$, where $o(1)$ denotes a function of k whose limit is 0. We conclude that the normalized boarding time satisfies $T(F, k) \le (\sum_i \sqrt{q_i})\sqrt{k} + o(1)$. We now show that $T(F, k) \ge ((\sum_i \sqrt{q_i}) - \varepsilon/\sqrt{q_i})\sqrt{k}$ for any $\varepsilon > 0$, To obtain the lower bound we fix any $u < 1$. Consider the horizontal line segment $U_{i,\varepsilon}$, consisting of points of the form $(r_{\sigma(i)}, q)$, with $s_i \le q < s_{i+1} - \varepsilon$. If $k > 1/varepsilon$, then the line segment joining the endpoint $(s_{i+1} - \varepsilon, r_{\sigma_i})$ of U_i to the initial point $(s_{i+1}, r_{\sigma(i+1)})$ of U_{i+1} has a positive slope if $\sigma_{i+1} > \sigma_i$ and a slope greater or equal $-k$ otherwise. In both cases it satisfies (4). The length of $U_{i,\varepsilon}$ is by (6 equal to $\sqrt{k}\sqrt{1/q_i}(q_i - \varepsilon)$, as required.

8 Summary and future work

We have introduced a multi-parameter tasks with precedences model, which captures the essential features of the airplane boarding process. The model allows us to compare various boarding policies. It also allows us to compute closed-form estimates for the boarding time in many cases of interest. The resulting model based computations have been validated against very detailed simulations of the boarding process of Van Landeghem and Beuselinck (2002). Both methods rank policies very similarly.

For an airplane with 6 passengers per row, the main candidates for a good boarding policy have been reduced to $F_{3,1}$ with seat type classification for 3 group policies and $F_{3,2}$ with seat type classification for 6 group policies. It is possible, using the analytic model, to reevaluate the effect of policy changes when the distance between rows or the number of passengers per row/per aisle change. Upon adding passengers per row or squeezing the distance between rows, the boarding process becomes slower due to congestion, and the F_1 policy more attractive among row based policies.

The most immediate challenge in terms of the analysis of airplane boarding seems to be the analysis of policies in which seats are unassigned. Based on sporadic personal evidence, we believe that such policies may in fact prove to be more efficient. The basic difficulty is in establishing a reasonable passenger behavior model which will lead to the construction of an appropriate probability distribution p. This will require extensive field work and interviews with customers. One insight which may be gained from the present work is that the congestion controls to a large extent boarding time. We believe that passengers which are strangers are naturally averse to congestion and will tend to space themselves along the airplane, thus effectively lowering the value of k. Other trends, such as a preference for front seats, have detrimental effects. However, until a more detailed study of passenger behavior is performed, we cannot provide a useful model for unassigned seating policies.

Acknowledgments

We wish to thank Percy Deift and Ofer Zeitouni for several helpful conversations and much encouragement. We would also like to thank Aharon Davidson, Matt Visser, Menkes Van den Briel and Jinho Baik for illuminating exchanges and discussions.

References

Bachmat, E., D. Berend, L. Sapir, S. Skiena. 2005. Airplane boarding, disk scheduling and space-time geometry, *Proceedings of the First Conference on Algorithmic Applications in Management (AAIM 2005)*, Lecture Notes in Computer Science Series vol. 3521, edited by N. Megido, Y. Xu and B. Zhu, published by Springer Verlag, 192–202.

Bachmat, E., D. Berend, L. Sapir, S. Skiena, N. Stolyarov. 2006. Analysis of aeroplane boarding via spacetime geometry and random matrix theory, *Journal of physics A: Mathematical and General* 39, L453–L459.

Ferrari, P., K. Nagel. 2005. Robustness of efficient passenger boarding in airplanes, *Transportation Research Board Annual Meeting*, paper number 05-0405. Washington D.C., Also `http://www.trb.org/`.

Ferrari P. 2005. Improving passenger boarding in airplanes using computer simulations, *International airport review.*

Marelli, S., G. Mattocks, R. Merry. 1998. The role of computer simulation in reducing airplane turn time. *Boeing Aero Magazine*, Issue 1.

Penrose R. 1972. Techniques of differential topology in relativity. *Regional Conference Series in Applied Mathematics 7*, published by SIAM.

Van den Briel, M., J. Villalobos, G. Hogg. 2003. The aircraft boarding problem. *Proc. of the 12'th Industrial Eng. Res. Conf., IERC*, CD ROM, article number 2153.

Van den Briel, M., J. Villalobos, G. Hogg, T. Lindemann, A.V. Mule. 2005. America West develops efficient boarding strategies, *Interfaces*, to appear.

Van Landeghem, H., A. Beuselinck. 2002. Reducing passenger boarding time in airplanes: A simulation approach. *European J. of Operations Research* 142, 294–308.

Appendix A: Table of permutations

We provide a table of the permutations which were used by Van Landeghem and Beuselinck in their experiments. As an example, the policy F_{6,σ_4}. The policy divides the passengers into six equal size groups. Assuming for simplicity of presentation that the airplane has 24 rows, then the order of boarding is:

rows 21-24,

followed by rows 9-12,

followed by rows 17-20,

followed by rows 5-8,

followed by rows 13-17,

followed by rows 1-4.

As an example of a half row class policy consider $F_{2,4,\sigma_{11}}$. The boarding order in this case is:

left side of the aisle, rows 19-24,

followed by right side of the aisle, rows 13-18,

followed by left side of the aisle, rows 7-12,

followed by right side of the aisle, rows 1-6,

followed by right side of the aisle, rows 19-24,

followed by left side of the aisle, rows 13-18,

followed by right side of the aisle, rows 7-12,

followed by left side of the aisle, rows 1-6.

The rest of the table can be understood in a similar manner.

0	2, 3, 1
1	4, 2, 3, 1
2	4, 1, 3, 2
3	6, 4, 2, 5, 3, 1
4	6, 3, 5, 2, 4, 1
5	6, 2, 5, 1, 4, 3
6	10, 8, 6, 4, 2, 9, 7, 5, 3, 1
7	10, 5, 9, 4, 8, 3, 7, 2, 6, 1
8	4, 1, 2, 3
9	6, 4, 5, 3, 1, 2
10	8, 6, 7, 5, 4, 2, 3, 1
11	8, 3, 6, 1, 4, 7, 2, 5
12	12, 10, 8, 11, 9, 7, 6, 4, 2, 5, 3, 1
13	12, 9, 11, 8, 10, 7, 6, 3, 5, 2, 4, 1
14	12, 4, 8, 5, 9, 1, 11, 3, 7, 6, 10, 2

Table 2: Permutations

Appendix B: Methods of computation

We provide explanations for the various calculations which were used in the computation of the table in Appendix B. We first recall the main formulas needed for the computations.

A) The basic formula for F_1 with $k > \ln 2$:

$$T(F_1, k) = \sqrt{k} + \frac{1 - \ln 2}{\sqrt{k}}\,. \tag{40}$$

B) The formula for F_2:

$$T(F_2, k) = \begin{cases} \sqrt{\frac{1}{2k}}(k + \frac{e^k - 1}{4}), & 1 \le k \le 2\ln 2, \\ \sqrt{2k} + \frac{3/4 - 2\ln 2}{\sqrt{2k}}, & k \ge 2\ln 2. \end{cases} \tag{41}$$

C) The formula for F_m, $m > 2$ and $k \ge 3/4 + \ln 2$:

$$T(F_m, k) = \sqrt{mk} - \frac{m-2}{\sqrt{mk}}(\ln 2 + 1/4) - \frac{2\ln 2 - 3/4}{\sqrt{mk}}\,. \tag{42}$$

D) The formula for $F_{m,\sigma}$, where $k \geq 3/4 + \ln 2 + \lambda(\sigma) - 1$:

$$T(F_m, \sigma, k) = \sum_{j=1}^{r(\sigma)} T(F_{b_j}, k) \frac{\sqrt{b_j}}{\sqrt{m}} - e \frac{1}{\sqrt{km}} \,. \tag{43}$$

E) The formula for $F_{2,m,\sigma}$:

$$T(2, m, \sigma, k)\sqrt{n} = \sqrt{2} T(m, \sigma, k/2) \,. \tag{44}$$

All policies not involving a permutation σ are computed directly from these formulas by setting $k = 4$ and dividing by the expression for F_1, which yields the value 2.15. We now consider the policies which involve a permutation σ.

$\sigma_0, \sigma_1, \ldots, \sigma_6$: We can apply (43) in all these cases. For σ_4 and σ_5 we use the fact that for $m = 2$ the argument leading to (43) only requires $k \geq \lambda(\sigma)$.

σ_7: In this case (43) does not hold since some of the U_i's are empty. To compute a maximal curve U, we need to decide which of the U_i's are non-empty. This can be done using a dynamic programming approach, which is easily carried out by hand for small cases. The maximal curve is obtained when U_i is non-empty for all even values of i. We note that, for even i, the top right corner of $S_{i,\sigma}$ blocks the bottom right corner of $S_{i+2,\sigma}$. As a result, $T = \frac{5}{\sqrt{10}} T(F_1, k)$ with $k = 4$.

$\sigma_8, \ldots, \sigma_{11}$: In these cases the performance is computed in a straightforward manner using the formulas above, in different combinations. In accordance with formula (44) we set $k/2 = 2$ in formulas (40)–(43).

σ_{12}: We have $r(\sigma) = 4$, $b_j = 3$, and all the U_i's are non-empty, but we cannot use (43) since $m = b_j > 2$ and the condition on k is not satisfied with $k/2 = 2$. The computations are made assuming $\delta_i = 1/4$ for $i = 3, 6, 9, 12$. It can be shown that this choice leads to a curve length which is at most 1 percent below the maximal length.

σ_{13}: As in the case of σ_7, the maximal curve is obtained when U_i is non-empty for even i. We have $k/2 = 2$, which yields $T = \frac{6}{\sqrt{12}} T(F_1, 2)$.

σ_{14}: By aisle symmetry, this involves the same calculation as σ_{12}.