Lorentzian-geometry-based analysis of airplane boarding policies highlights "slow passengers first" as better

Sveinung Erland[*]

Department of Maritime Studies, Western Norway University of Applied Sciences, 5528 Haugesund, Norway

Jevgenijs Kaupužs

*Faculty of Materials Science and Applied Chemistry, Institute of Technical Physics, Riga Technical University, 1048 Riga, Latvia
and Institute of Mathematical Sciences and Information Technologies, University of Liepaja, 3401 Liepaja, Latvia*

Vidar Frette

Department of Fire Safety and HSE Engineering, Western Norway University of Applied Sciences, 5528 Haugesund, Norway

Rami Pugatch

Department of Industrial Engineering and Management, Ben-Gurion University, Beer-Sheva 84105, Israel

Eitan Bachmat

Department of Computer Science, Ben-Gurion University, Beer-Sheva 84105, Israel

We study airplane boarding in the limit of a large number of passengers using geometric optics in a Lorentzian metric. The airplane boarding problem is naturally embedded in a $(1 + 1)$-dimensional space-time with a flat Lorentzian metric. The duration of the boarding process can be calculated based on a representation of the one-dimensional queue of passengers attempting to reach their seats in a two-dimensional space-time diagram. The ability of a passenger to delay other passengers depends on their queue positions and row designations. This is equivalent to the causal relationship between two events in space-time, whereas two passengers are timelike separated if one is blocking the other and spacelike if both can be seated simultaneously. Geodesics in this geometry can be utilized to compute the asymptotic boarding time, since space-time geometry is the many-particle (passengers) limit of airplane boarding. Our approach naturally leads to the introduction of an effective refractive index that enables an analytical calculation of the average boarding time for groups of passengers with different aisle-clearing time distribution. In the past, airline companies attempted to shorten the boarding times by trying boarding policies that allow either slow or fast passengers to board first. Our analytical calculations, backed by discrete-event simulations, support the counterintuitive result that the total boarding time is shorter with the slow passengers boarding before the fast passengers. This is a universal result, valid for any combination of the parameters that characterize the problem: the percentage of slow passengers, the ratio between aisle-clearing times of the fast and the slow group, and the density of passengers along the aisle. We find an improvement of up to 28% compared with the fast-first boarding policy. Our approach opens up the possibility to unify numerous simulation-based case studies under one framework.

I. INTRODUCTION

A main theme in statistical physics is the connection between the microscopic dynamics of an ensemble of interacting particles or units and the macroscopic observables. In this paper we consider the problem of airplane boarding, where the microscopic level is the structure of the passenger queue and the main macroscopic observable is the required time for all N passengers to get settled in their assigned seats. The connection between the microscopic level and the macroscopic observable is presented by a simple two-dimensional diagram, with a geometric interpretation that is directly linked to special relativity.

The average boarding time can be calculated analytically in the large-N limit. This average boarding time has been found to be a square-root law in the number of passengers [1–8]. Our analytical result further enables a direct approach for optimization over the three main parameters of the boarding process, to be elucidated in what follows. Interestingly, our approach presents a very simple and straightforward interpretation to the causal set program of quantum gravity [9–12]. In

----

[*]Author to whom correspondence should be addressed; sver@hvl.no

terms of the causal set program, the main contribution of this work is that we consider a scenario in which there is more than one type of space-time event, with different types of events having different proper time contributions.

During airplane boarding, the passengers have reserved seats, but arrive at the gate in arbitrary order, and a queue of passengers is formed in the jet bridge. Recently, it has been shown that any delay in the boarding process immediately adds to the overall airplane turnaround time, especially for intracontinental flights, i.e., delays in the boarding process will delay the flight departure time [13,14]. Nevertheless, attempts to optimize the boarding process have not been so successful, as many travelers still experience slow queue advancement, with a few busy passengers arranging carry-on luggage and taking their seat at any given moment during boarding.

In fact, it can be easy to arrange a boarding queue optimally (an example is given in Appendix C). Still, optimal queue arrangements can hardly be used, for two reasons. First, even though a specific queue position for each passenger may be implemented, such instructions tend to reduce passenger satisfaction, for example, through splitting groups of passengers traveling together. Second, the optimal queue arrangements are usually not robust to deviations in terms of passengers who do not take their designated position in the queue. So, unlike many other problems, even though optimal solutions could be easily found, the optimal solutions for airplane boarding are of little interest for the airlines due to their impracticality.

Weaker constraints, for example, with passengers assigned to groups according to seat or row number, are employed more often. We will refer to an imposed arrangement of the queue prior to boarding as a boarding policy. The most common policy is the unorganized random boarding policy, where the passengers enter the queue in random order. Surprisingly, random boarding is relatively efficient. Another familiar policy is the back-to-front policy, where the passengers are divided into two or more groups and those who have designated seats in the back of the airplane are instructed to enter first. This is a widely used policy, but both simulations and analytical computations show that it is usually detrimental compared to random boarding [4].

In this paper we investigate two simple group-based policies that can bring us closer to a near-optimal solution, which is nevertheless practical, namely, passengers who tend to use more time to complete the seating are separated from the others and can be requested to enter the airplane either before or after the remaining passengers. One such type of passengers are those with overhead bin luggage.

Moreover, the size of carry-on luggage is also an easy and practical criterion for separating the passengers into groups of who we denote slow and fast passengers, respectively. Fast-first boarding policies, where passengers without overhead bin luggage are allowed to board before other passengers enter the plane, have been implemented in the past [15]. The opposite slow-first policy has been implemented in the way that certain small groups of slow passengers usually are allowed to board before other passengers. Such groups typically consist of, e.g., small children and those who need special assistance.

That the slow-first policy is superior to the fast-first policy is surprising at first. We apply the analytical tools of Lorentzian geometry to prove that this is indeed the case. While Lorentzian geometry has been used previously to analyze the boarding process [1,7], the two-group scenario introduces an additional aspect. The different group speeds can be analyzed in terms of the concept of a refraction index, i.e., they scale the metric by different amounts in different regions of space, introducing an optics perspective into the picture.

Other techniques in the airplane boarding literature are not able to analyze such boarding policies in a general setting (see, e.g., [13]). Simulation algorithms may indicate which policy is superior to others, but fail to provide generality or proof.

To be more specific, the simulation-based result in [16] is here verified for far more general model settings. Policies that organize the fast and slow passengers in more optimal ways according to their designated row positions have been constructed [17–20]. However, as mentioned before, policies that require detailed control on each of the passengers cause great discomfort and noncompliance. Thus, while these boarding policies are of theoretical interest, they are not expected to be implemented.

The structure of the paper is as follows. The boarding process is described in Sec. II and the main results are summarized in Sec. III. Relevant parameters of the boarding process, its representation in diagrams, and the connection to space-time geometry are presented in Sec. IV. The asymptotic boarding time is presented in Sec. V for general group-based policies in the many-passenger limit ($N \rightarrow \infty$). In Sec. VI we compute analytically the asymptotic boarding time for the slow-first (SF) and the fast-first (FF) policies. We also show by simulations that the large-N limit results hold for a realistic number of passengers.

II. BOARDING PROCESS

We consider the boarding process from the time when the passengers have queued up in the jet bridge outside the airplane entrance until the last passenger is seated. Most passengers wait most of the time during boarding since they are blocked by the other passengers from reaching their designated seat. We assume that the queue order is maintained throughout the process, i.e., passengers cannot pass other passengers that are in front of them in the aisle.

The boarding is modeled as an iterative two-step process. First, all passengers move until they either reach their designated row or until they are blocked on the way to their seat by other passengers. This is assumed to take an insignificant amount of time compared to the next step, where those passengers who stand next to their designated row use a certain aisle-clearing time in order to organize luggage and take a seat.

A simple example with only $N = 8$ passengers, all with the same aisle-clearing time, is presented in Fig. 1. At each time step, a group of passengers is able to sit down simultaneously. A passenger can be delayed by the passenger in front in two ways. First, the passenger could have a higher row number than the passenger in front (in time step $t = 1$, the third passenger who is heading for row 4 must wait for the passenger taking a seat at row 2). Second, the passenger could be displaced by passengers who are waiting for other passengers
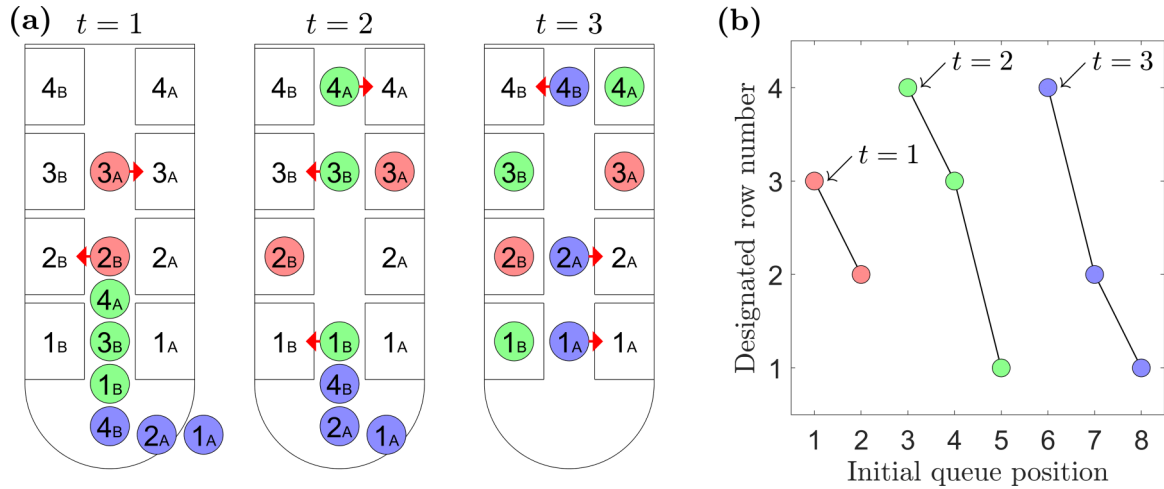
FIG. 1. (a) Illustration of the stepwise advance of the queue during the boarding process, with $N = 8$ passengers, four rows and two seats per row. There is space for two passengers per row along the aisle. Each passenger is marked as a circle with designated row number. At each time step, the queue moves forward, and passengers that have reached their designated rows sit down simultaneously. Red arrows indicate passengers that take their seat at that time step, and each group of passengers that sit down simultaneously is color coded. (b) Each passenger is marked as a point in a qr diagram: The initial queue position of the passenger is on the horizontal axis and the designated row number is on the vertical axis. The points of passengers that sit down simultaneously are joined by line segments.

to take a seat (in time step $t = 1$, the fifth passenger is heading for row 1, but must wait until passengers in front have been let through). This displacement effect is less significant if passengers stand closer to each other and thus occupy less space in the aisle. We call the time until the last passenger is seated the boarding time T. In Fig. 1, $T = 3$ time steps.

By reorganizing the queue in Fig. 1, it is possible to obtain a minimal boarding time, as shown in Appendix C. For such optimal solutions, it is necessary to impose a specific position in the queue for each passenger. Optimization at the level of individual passengers will not be pursued in this article.

For the sake of visualizing and analyzing the boarding process, we present the qr diagram in Fig. 1(b). In the qr diagram a point (q, r) represents a particular passenger's initial queue position q and designated row number r in the airplane. Passengers that take seats simultaneously are linked by lines. We call each such group of passengers a wave front in analogy to wave fronts in physics, as they represent all the events that share the same phase, i.e., all the passengers that are seated simultaneously. The boarding time can be found by counting the number of equidistant wave fronts, multiplying by the time difference.

We use qr diagrams as a tool to analyze the boarding process. Such diagrams convey the entire hierarchy of blocking between passengers for a given queue configuration.

III. MAIN RESULTS

The main results of this paper are shown in Fig. 2 for a particular realization of the three governing parameters to be defined below. In the subsequent sections we will prove rigorously that the main features in Fig. 2 are universal and apply for any set of parameters.

The airplane boarding problem we consider here is characterized by three key parameters. The first parameter is the congestion k, which is the ratio between the queue length
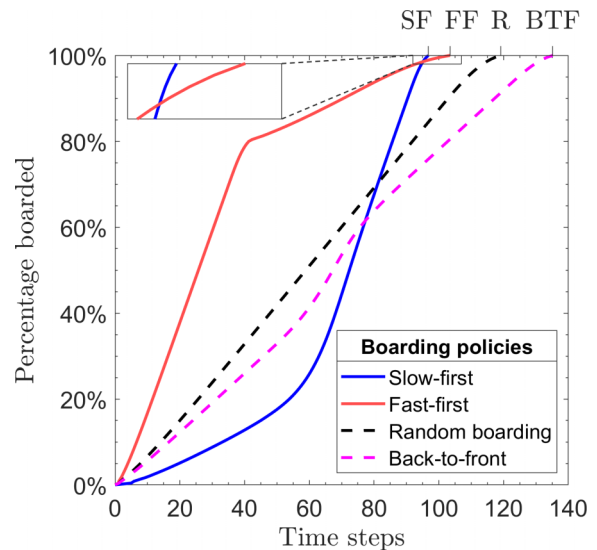

FIG. 2. Comparison of four different boarding policies. We used a realistic congestion parameter $k = 4$, 20% slow passengers ($p = 0.2$), and assumed that the slow passengers clear the aisle on average five times slower than the fast passengers ($C = 0.2$). We assumed there is a single aisle and 6 seats per row, and a total of $N = 240$ passengers. The percentage of seated passengers is plotted as a function of time. Remarkably, on average, the fast-first policy (leftmost, red solid line) is leading all the way to around ∼98%. However, the slow-first policy eventually seats all passengers in a shorter time, relative to all the other policies. The fast-first policy is second (FF, +7%), random boarding comes third (R, +23%) and the back-to-front policy turns out to be the worst (BTF, +40%). That the slow-first policy is superior can be intuitively explained by that it is the most parallel among all the policies, i.e., it better exploits the possibility to seat passengers simultaneously, while the other policies are more serial in structure. The graph is an average of 10 000 discrete-event runs.

before boarding to the aisle length. For common passenger airplanes, values of k are typically in the range of 3–5. The second parameter is the fraction p of passengers who are considered slow, i.e., passengers with long aisle-clearing time. The remaining fraction $1 - p$ are considered fast passengers. The third parameter C is the ratio of the aisle-clearing time of the fast passengers to the aisle-clearing time of the slow passengers.

In Fig. 2, the congestion parameter is $k = 4$, the fraction of slow passengers $p = 20\%$, and the ratio of the aisle-clearing time of the fast passengers to that of the slow passengers $C = 0.2$, i.e., the aisle-clearing time of the slow passengers is five times longer than that of the fast passengers.

The graphs in Fig. 2 show the percentage of seated passengers as a function of time for the four different boarding policies described in Sec. I, namely, slow passengers first, fast passengers first, back-to-front boarding, and random boarding. The boarding is completed when the fraction of seated passengers reaches 100%. The boarding time T is equal to the time when all the passengers are seated. In the figure, the following ranking of the policies can be observed. The best boarding policy is slow first, with a boarding time of 97 time steps; the second fastest boarding policy is fast first, with a boarding time of 103; the third best policy is random boarding with a boarding time of 119; and the worst policy is back to front with a boarding time of 135. In Ref. [4] it is shown that random boarding is typically superior to back to front. Random boarding and back to front are included in this figure for reference only, since they have been studied in the literature and are often implemented by airlines. Our focus remains the comparison between slow first and fast first.

The graphs for SF and FF in Fig. 2 both consist of two curve segments with different slopes. For fast first, the steep segment comes first, followed by the less steep, with opposite order for slow first. The steep segments correspond to boarding dominated by fast passengers, while the less steep segments are dominated by slow passengers. As boarding starts, the queue of passengers is four times as long as the aisle ($k = 4$). Thus, in the fast-first case, the first slow passenger arrives in the aisle late during the boarding process. Similarly, for slow first, only after a significant portion of the slow passengers are seated, the first fast passengers enter the airplane.

Comparing the graphs of fast first and slow first in Fig. 2 more closely, it is clear that while the curves are quite similar during the fast and the slow regimes, the transitions between the regimes are different. The trajectory of the fast-first policy has a distinct change in slope around time step 40. The corresponding transition in the slow-first policy is much smoother. This indicates that a significant proportion of the fast passengers are able to take their seat simultaneously with the last slow passengers. This is not the case with the fast-first policy, since only a few slow passengers are able to take their seat during the relatively short time period that it takes the last few fast passengers to sit. This asymmetry explains why the slow-first policy is superior.

For all four policies in Fig. 2, the boarding time is determined by the longest chain of passengers, similar to the ones shown in the qr diagrams in Fig. 3. The preceding passenger in a chain must take one's seat before the next in the chain can sit down. The boarding time is the sum of the aisle-

clearing times for passengers that belong to the longest chain. The longest chain follows the asymptotic longest curve (the geodesic), up to statistical fluctuations that are diminishing as the number of passengers increases ($N \to \infty$). For slow first and fast first in Figs. 3(c) and 3(d), respectively, the aisle-clearing time is twice as long for passengers in the slow group than for those in the fast group ($C = 0.5$). Note that the aisle-clearing time acts as a refraction index as the longest curve is refracted at the border between the two groups.

Figure 2 reports boarding times for one specific choice of the parameters k, p, and C. Under variations in these parameters, the comparison can conveniently be made using the relative difference

$$D(k, p, C, N) = \frac{\langle T_{\text{FF}} \rangle - \langle T_{\text{SF}} \rangle}{\langle T_{\text{SF}} \rangle}$$

between the average boarding times of the fast-first and slow-first policies. The contour plot in Fig. 4(a) shows D in the (p, C) unit square for $k = 4$ in the asymptotic case when $N \to \infty$. It is obtained by the space-time geometry approach. The average boarding time is larger for fast first than for slow first for all values of p and C. For $(k, p, C) = (4, 0.2, 0.2)$, the relative difference is $D = 11\%$ for $N \to \infty$, compared to $D = 7\%$ for $N = 240$ in Fig. 2. The maximum of $D = 20\%$ (for $k = 4$) is obtained for (p, C) very small.

Our main result can be stated as follows: The expected boarding time $\langle T \rangle$ is shorter for the slow-first policy than for the fast-first policy for all values of $k > 0$ and $p, C \in (0, 1)$, in the asymptotic regime when $N \to \infty$. The maximum relative difference between the policies when $N \to \infty$ is at least $D = 28.4\%$. This result is proved analytically in Appendix B. The maximum relative difference is $D = 28.4\%$ for $k = 1.594$, $C = 0.513\sqrt{p}$, and p small. Even when the fraction of slow passengers is fixed to the more realistic value $p = 0.1$, the maximum relative difference is $D = 24.4\%$ (for $k = 1.54$ and $C = 0.16$).

In Fig. 4(b) the relative difference in boarding times for $k = 4$ and finite N are compared with the asymptotic results when $N \to \infty$. All simulation results show $D \geqslant 0$. The relative difference is larger in the $N \to \infty$ limit compared with finite-N cases, but still the relative ranking between the results for different parameter settings is to a large degree preserved for smaller values of N.

IV. BOARDING PROCESS AND SPACE-TIME GEOMETRY

In this section we explain further the analogy between airplane boarding and space-time geometry. The reader is referred to Ref. [7] for a more rigorous mathematical description.

During boarding, a given passenger may be blocked from reaching one's designated row by another passenger, which in turn may be blocked by others. This blocking hierarchy can be visualized in the qr diagrams through so-called blocking chains. Importantly, the longest blocking chain determines the boarding time.

A condition for blocking to occur is formulated in Eq. (1) (Sec. IV C). The condition is extended to the continuous case when $N \to \infty$ and then the passengers correspond to events in space-time geometry. In this setting, the blocking
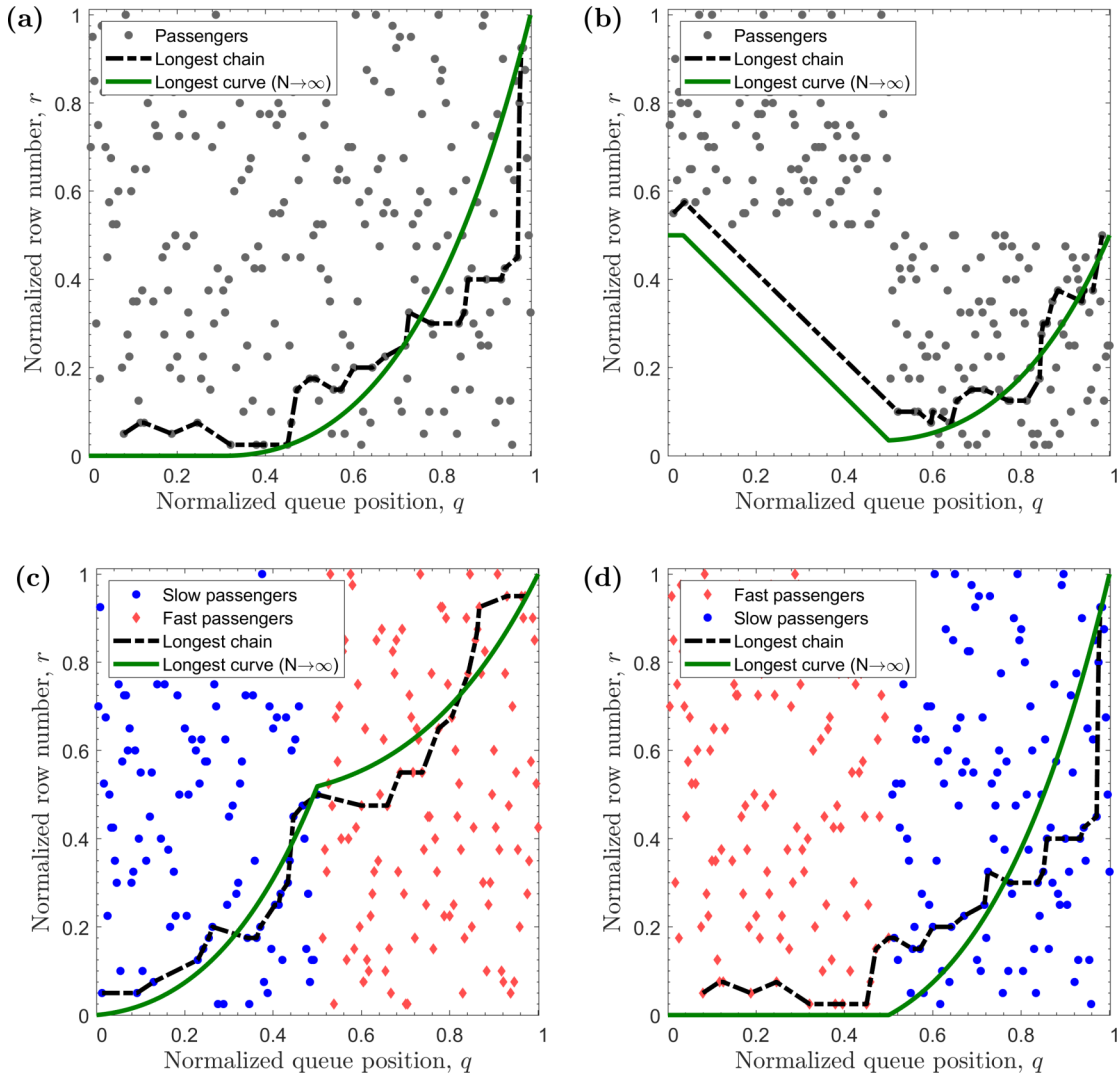
FIG. 3. The qr diagrams for four different boarding policies, with each of the $N = 240$ passengers marked as a point, $h = 6$ seats per row, and congestion $k = 1$. (a) Random boarding policy: The passengers are uniformly distributed over the diagram. (b) Back-to-front policy with two equal-sized groups: The first part of the queue is heading for the rows in the back of the airplane. (c) Slow-first policy with two equal-sized groups: The slow passengers are in the first part of the queue (blue bullets). (d) Fast-first policy with two equal-sized groups: The fast passengers are in the first part of the queue (red diamonds). For all policies the boarding time is the sum of the aisle-clearing times for passengers that belong to the longest chain (dashed lines). The preceding passenger in a chain must take his seat before the next in the chain can sit down. In all four diagrams, the longest chain that determines the boarding time follows the asymptotic limit (solid line, the geodesic), up to statistical fluctuations that are diminishing as the number of passengers increases (see Sec. VI for further details).

condition also determines, up to a proportionality constant, the appropriate Lorentzian metric that should be used to calculate the distance (proper time) along a trajectory between two events (Sec. IV D). Finally, the boarding time can be found by computing the longest blocking chain (Sec. IV E) which, in the limit $N \to \infty$, tends to the length of the geodesic line.

A. Main parameters

The boarding process is governed by the following parameters.

(i) The total number of passengers N. For simplicity, we assume that there are no empty seats in the airplane, i.e., the

airplane is full. Hence, the total number of passengers equals the number of seats in the airplane. In Fig. 1, $N = 8$.

(ii) Congestion k, the length of the queue before boarding $(t = 0)$ *relative to the length of the aisle.* Let h be the number of seats per row, w the distance between passengers needed for each to stand comfortably, one after the other along the aisle, and d the distance between consecutive rows. Then $k = hw/d$. The parameter k reflects the interior design of the airplane and the maximum density of passengers queuing along the aisle.[1] In Fig. 1, the value $k = 1$ is used for simplicity.

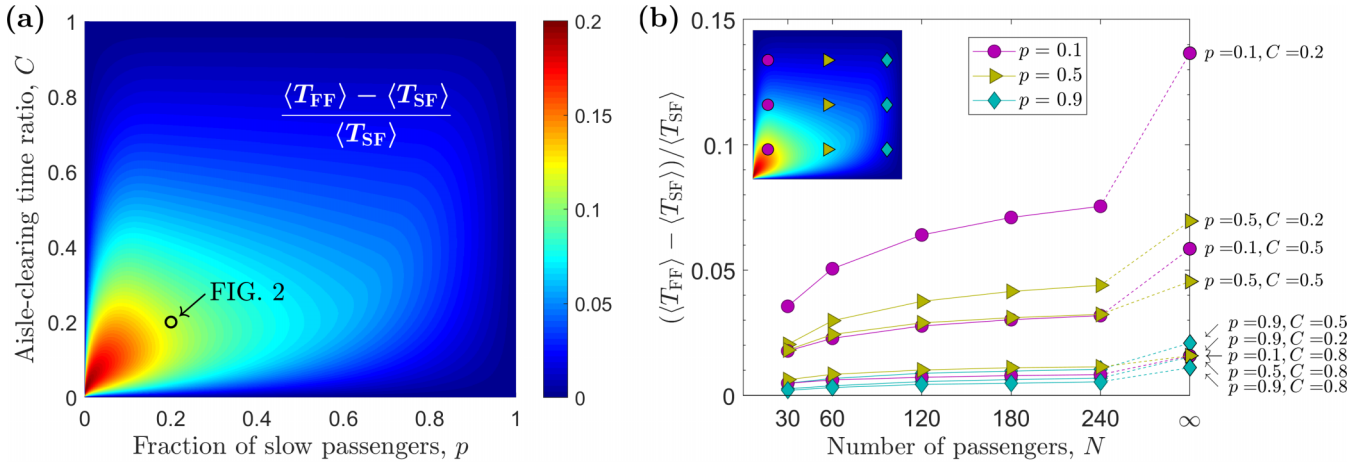[1]The parameter k can also be modified to include the number of aisles and the relative occupancy of the airplane [21].

FIG. 4. Relative difference in average boarding time $D = (\langle T_{\text{FF}} \rangle - \langle T_{\text{SF}} \rangle)/\langle T_{\text{SF}} \rangle$ between the fast-first and the slow-first policies when $k = 4$. (a) Number of passengers $N \to \infty$. The slow-first policy is superior for all (p, C) combinations, and the maximum relative difference is 20% for small p and C. For the parameter choice in Fig. 2, the relative difference is 11% (black circle). (b) Simulation results for finite numbers of passenger N confirm that slow-first is superior to fast-first, i.e., $D = (\langle T_{\text{FF}} \rangle - \langle T_{\text{SF}} \rangle)/\langle T_{\text{SF}} \rangle > 0$ for increasing N for all combinations of parameter values $p \in \{0.1, 0.5, 0.9\}$ and $C \in \{0.2, 0.5, 0.8\}$. The rightmost points are asymptotic values taken from the indicated positions in the inset contour plot from (a). There are six seats per row, $k = 4$, and the accuracy is ± 0.0002 (as a result of 10^6 runs for each finite-N data point).

(iii) Fraction of slow passengers p. In Fig. 1, all passengers have equal aisle-clearing time, so $p = 0$.

(iv) Aisle-clearing time τ, the time needed for a passenger to organize bin luggage and take a seat. In Fig. 1 all passengers have an aisle-clearing time of $\tau = 1$ time steps.

(v) Queue position of a passenger q normalized by the total number of passengers N. In Fig. 1, the fourth passenger in the queue at $t = 1$ (aiming for row 3) has $q = \frac{4}{8} = 0.5$.

(vi) Designated row number for a passenger r normalized by the total number of rows. The fourth passenger in the queue in Fig. 1 has $r = \frac{3}{4} = 0.75$.

In Fig. 1 the actual queue and row numbers are given on the axes, while in the rest of the paper the normalized (q, r) values will be used.

B. Boarding policies visualized

A boarding policy is the way the queue of passengers is organized. The most common policy is the unorganized random boarding policy, where the passengers enter the queue in random order. A typical scenario with the random boarding policy is illustrated in the qr diagram in Fig. 3(a). The points representing each of the $N = 240$ passengers are uniformly distributed over the unit square.[2]

A scenario with the back-to-front policy is shown in Fig. 3(b). The passengers are divided into two groups, where those who have designated seats in the back of the airplane constitute the first part of the queue. Within each group,

the passengers are randomly distributed in the queue. In Fig. 3(c) a scenario with the slow-first policy is shown. The diagram resembles the one of the random boarding policy, but passengers assumed to use a long time to take a seat are placed in a separate group in the first part of the queue. The designated row numbers are randomly distributed within both groups as in random boarding. The fast-first policy in Fig. 3(d) has a diagram similar to that of the slow-first policy, except that the fast passengers here appear in the left part of the diagram.

C. Blocking chains and blocking relation

As shown in Fig. 1, the boarding process can be thought of as wave fronts of passengers that take their seats simultaneously. When all passengers that stand next to their designated row have taken their seat, the remaining passengers in the aisle move rapidly forward and a new wave front of passengers sits down. Hence, the boarding time is the product of the aisle-clearing time times the number of wave fronts needed to seat all passengers. Wave fronts are shown in Fig. 1, with $N = 8$ passengers and congestion parameter $k = 1$. When $k = 4$, as in the random boarding case in Fig. 5(a), the wave fronts are steeper and a single wave front spans less of the q axis since there is no room for more than a quarter of the initial queue in the aisle.

The direct approach of finding all the wave fronts and counting them to determine the boarding time is impractical when the number of passengers N is large. Furthermore, the average boarding time, obtained when averaging over all possible queue configurations, is even harder and closed-form analytical results for finite N are not known. In the following we will describe an indirect way to calculate the number of wave fronts (the boarding time) by the introduction of blocking chains, which later will be shown to correspond to

---

[2]This uniformity applies in general for a coarse-grained description when N is sufficiently large. On the microscopic level, however, the point cloud has a structure since the q and r directions are not equivalent: For a given q value one has only one point in the diagram, while for a given r value there are as many points as there are seats in a row.
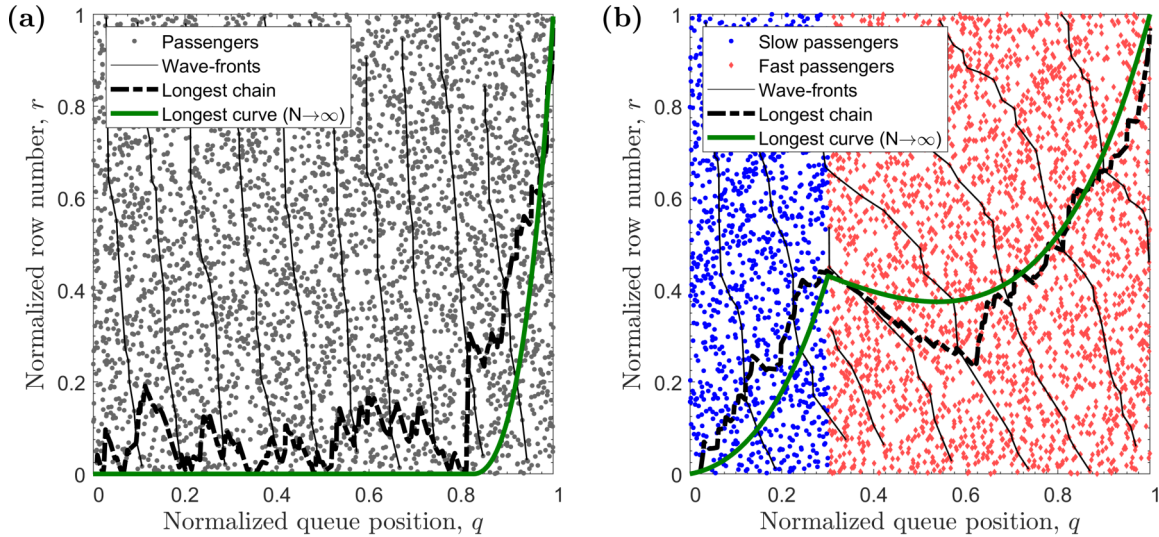
FIG. 5. With $N = 4000$ passengers, there is closer correspondence between the longest chain and the longest curve than in Fig. 3 (where $N = 240$). Passengers (points) seat simultaneously in consecutive wave fronts (black thin lines). For clarity, only every 20th wave front is shown. (a) Random boarding policy: $k = 4$, with the same aisle-clearing time for all passengers. (b) Slow-first policy: $(k, p, C) = (1.5, 0.3, 0.33)$. The longest curve is refracted when the aisle-clearing time changes value on the border between the slow (blue bullets) and the fast (red diamonds) passengers. The aisle-clearing time plays the role of a refractive index.

causal chains in space-time geometry. These will be essential in order to establish the asymptotic boarding time as $N \rightarrow \infty$.

We say that passenger A blocks passenger B if A must be seated before B can sit down. Two passengers in the same wave front cannot block each other and A cannot block B if B is in front of A in the queue. A blocking chain consists of passengers that consecutively block each other. The length of the chain is the sum of aisle-clearing times for passengers that belong to the chain, and the length of the longest chain among all blocking chains equals the boarding time. Given a queue, we can construct the longest chain by starting with one of the passengers in the last wave front. Several passengers in the preceding wave front may be blocking this passenger. The one that is closest in the queue is chosen as the next passenger in the chain. The longest chain is obtained by proceeding like this, until reaching a passenger in the first wave. Examples of the longest chains are shown in Figs. 3 and 5. Notice that the longest chains are approaching the asymptotic longest curve when N increases (see details in the following sections).

The blocking chain can be defined in terms of a blocking relation. Let passenger A be in front of passengers B_1 and B_2 in the queue, as shown in Fig. 6. Passenger B_1 is heading for a row farther back in the airplane, so A is obviously blocking B_1. Passenger B_2 is heading for row 1, which is in front of A's row. However, due to the displacement caused by the two passengers in between A and B_2, A blocks B_2 from reaching B_2's seat.[3]

More generally, let $g(A, B)$ be the number of passengers (including passenger A) standing in between passengers A and B, just before A sits down. Passenger A is in front of B and they are heading for rows R_A and R_B, respectively. We say passenger A blocks passenger B if the distance between the designated rows of A and B is less than the space (displacement) in

the aisle occupied by the passengers between A and B:

$$(R_A - R_B)d < g(A, B)w. \tag{1}$$

Here d is the distance between each row and w is the space (length) occupied by each passenger along the aisle.

In Fig. 6, where $d = 2w$, both B_1 and B_2 are blocked by A according to this definition. For B_1 the row distance is $(R_A - R_{B_1})d = -2d = -4w$ and the displacement is $g(A, B_1)w = w$. For B_2 the respective values are $(R_A - R_{B_2})d = d = 2w$ and $g(A, B_2)w = 3w$. This means that A blocks both B_1 and B_2, since the inequality in Eq. (1) is satisfied in both cases. In fact, most of the other passengers are blocked by A in this example. We call the inequality in Eq. (1) a blocking relation, and it turns out that the relation also determines how the number of passengers in a blocking chain can be computed when $N \rightarrow \infty$.

D. Blocking relation and curve length in space-time geometry ($N \rightarrow \infty$)

In relativity theory, events are given by space and time coordinates and the mathematical description is continuous. Events can be classified through the past-future (or causality) relation. Event A is in the past of an event B if it is possible to move from A to B under or at the speed of light. In other words, a future event B can only be affected by an event A if B is within the future light cone of A.

Proper time is the time passing on a clock attached to a particle passing through a (continuous) set of events. In the causal set approach to gravity, space-time is discrete and composed of a finite number of events, each contributing one time unit to proper time. In this approach, continuous space-time emerges as the limit of discrete space-time as the number of events increases.

For airplane boarding, the passengers play the role of the events. Passengers have a natural causal (past-future) structure

_____

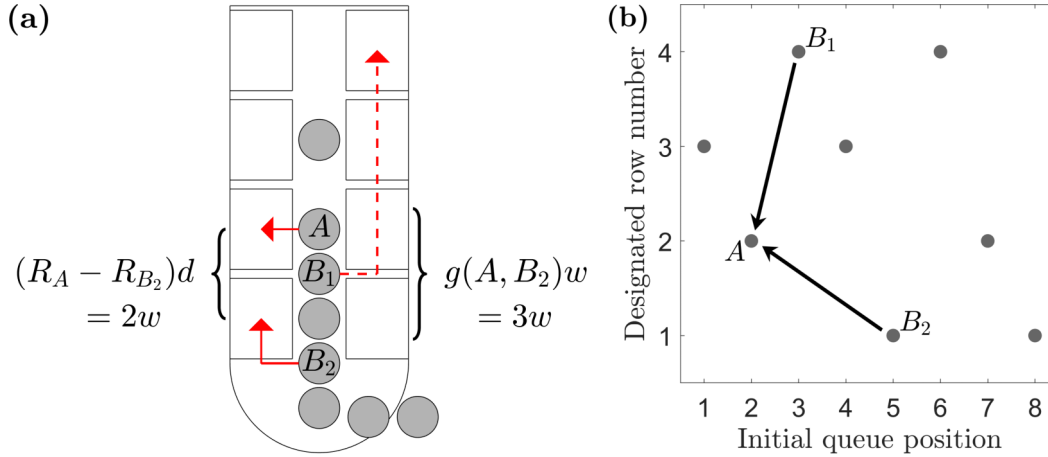[3]For $k = 0$, blocking through displacement never occurs.

FIG. 6. Example of blocking relations for the same case as in Fig. 1. The distance d between consecutive rows is twice the distance w occupied by each passenger in the aisle. Passenger A blocks both passengers B_1 and B_2. (a) Passenger B_1 is blocked by A since B_1 is heading for row 4, which is beyond row 2 where A is taking a seat. Passenger B_2 is blocked by A by displacement as in Eq. (1): The space $3w$ occupied in the aisle by passenger A and the two passengers between A and B_2 is larger than the distance $d = 2w$ between the designated rows of A and B_2. (b) The blocking relations are indicated by arrows. There are several other blocking relations that are not shown in the figure.

defined by the blocking relation in Eq. (1). We consider a passenger A as being in the past of passenger B if passenger B is blocked by A. In this setting, we may call the blocking chain a causal chain, where each passenger in the chain contributes its aisle-clearing time to the proper time of the chain. Thus, the proper time of the longest chain equals the boarding time.

The continuous analog of the notion of a causal chain is a causal curve, i.e., the possible trajectory of an object traveling below the speed of light. In space-time geometry the proper time of a particle's path (causal curve) is given by its length as found by integration using the space-time interval ds. In its simplest form with one spatial dimension, it is given by $ds^2 = dt^2 - dx^2$. The past-future relation, which ensures that future events stay within the future light cone of past events, is given by $ds^2 \geqslant 0$.

Using the coordinate transformations $q = t + x$ and $r = t - x$, the queue-row diagram emerges from the future light cone of a space-time diagram as shown in Fig. 7(a). The space-time interval now reads

$$ds^2 = dt^2 - dx^2 = (dt + dx)(dt - dx) = dq\,dr.$$

If passengers had no width ($w = 0$), the congestion parameter would be $k = hw/d = 0$, and the blocking relation in Eq. (1) reduces to $R_A < R_B$ when passenger A is in front of B in the queue. This means that passenger A in Fig. 7(a) can only block passengers in the shaded, upper right-hand rectangle of point A. Notice the past-future relation $ds^2 = dq\,dr \geqslant 0$, i.e., $dr/dq \geqslant 0$, is equivalent to the blocking relation; line segments between events (passengers) must be nondecreasing in q.

Real passengers are not infinitely thin and the congestion parameter $k = hw/d > 0$. Assume that N is large. Let passengers A and B be close in the queue, A in front of B, separated by $dq > 0$ [see Fig. 7(b)]. The normalized difference in assigned row position is $dr = (R_B - R_A)/(N/h)$, where h is the number of seats per row. Just before passenger A sits down, the number of passengers between A and B in the queue is essentially those who are heading for the rows behind

R_A. Since the passengers are uniformly distributed on the qr diagram, this number is given by $g(A, B) \approx dq(1 - r)N$, shown as the shaded area in Fig. 7(b). The blocking relation in Eq. (1) can now be written

$$\frac{h}{Nd}[(R_B - R_A)d + g(A, B)w] \approx dr + k(1 - r)dq > 0. \quad (2)$$

Thus, a causal curve must satisfy $r'(q) > -k(1 - r)$. It follows that the sectors of passengers blocked by A_1 and A_2 in Fig. 7(c) increase for smaller r; the potential for blocking other passengers is larger when the designated row is in the front of the airplane.

The proper time of a causal chain when each event in the chain contributes one time unit to the proper time and $N \to \infty$ corresponds to the proper time of a causal curve. There exists a Lorentzian metric ds that can be used to compute the length of a causal curve (its proper time) and this metric is defined uniquely, up to a constant scaling factor, by the blocking relation. For airplane boarding and the blocking relation in Eq. (2), the metric[4] is [7]

$$ds^2 = dq[dr + k(1 - r)dq]. \quad (3)$$

The length of a curve $r(q)$ between two points q_0 and q_1 is then given by

$$L(r) = \int_{q_0}^{q_1} \sqrt{r'(q) + k[1 - r(q)]}\,dq. \quad (4)$$

This definition of length, together with an appropriate scaling factor, will in the following sections be used to calculate the number of passengers in the longest chain in airplane boarding in the large-N limit.

———

[4]This is a simplified metric. The more general metric in [7], which includes the density distribution of passengers, has been used to analyze, e.g., the back-to-front boarding policy.
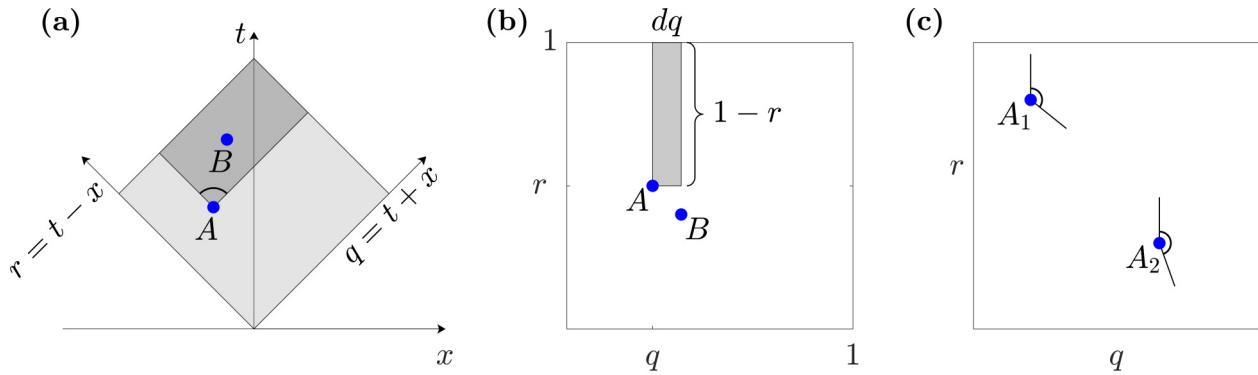
FIG. 7. (a) Coordinate transformation from space-time to queue row when $k = 0$. The qr diagram is in the future light cone of the origin. Passenger B can only be blocked by A if B is within the light cone of A. (b) Blocking by displacement for $k > 0$ according to Eq. (1). Since passenger B's designated row is in front of A's designated row, A can only block B by displacement. The value of k and the number of passengers in between A and B in the queue determine the extent of the displacement. Most of the passengers in between are heading for the rows behind A's designated row (shaded area). (c) Passengers nearby A_1 and A_2 must be within their respective future light cones in order to be blocked. The cones of each blocking passenger have a wider angle when the designated row is near the front of the airplane, reflecting a larger potential for blocking fellow passengers (here $k = 4$). Passengers can never block passengers who are in front of them in the queue, and therefore the line emanating upward is always part of the light-cone boundary.

E. Longest curves under the Lorentzian metric

According to general relativity theory, among all possible paths between two events, a free-falling particle (a particle only under the influence of gravity) will follow a trajectory that maximizes proper time (locally, between any two nearby points on the trajectory). Such free-fall trajectories are known as geodesics. The longest chains, which determine the boarding time, correspond to the longest curves in the continuous version. The longest curves are geodesics when not constrained by boundary conditions.

We first look at the Euclidean metric, where the length of a curve is $\int \sqrt{r'(q)^2 + 1} \, dq$. The equidistant points relative to a starting point in $(0, 0)$ are circles, as shown in Fig. 8(a). The shortest curves from the starting point to any other point are
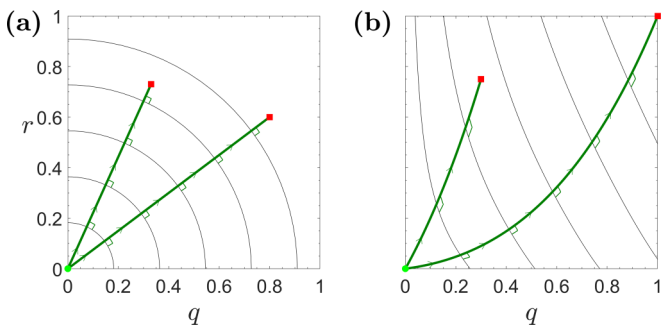

FIG. 8. The optimal path (green thick curves) depends on the metric and is orthogonal to the contour lines (black thin curves) which defines the equidistant points from the starting point (origin). (a) Under the Euclidean metric, the contour lines are circle shaped. The shortest paths from the starting point to any other point (red squares) are straight lines. (b) An appropriate Lorentzian metric is used in airplane boarding, and the starting point is $(0, 0)$ in the qr diagram (here $k = 0.6$). The contour lines coincide with passenger wave fronts when $N \to \infty$. The longest curves (geodesics) from the starting point to any other point are everywhere orthogonal to the contour lines (wave fronts).

straight lines that are orthogonal to the circles. This can be compared to a ball rolling down a hill where the contour lines indicate the height. The shortest path is always in the steepest direction, orthogonal to the contour lines, and the ball will take the same time to reach any point on a chosen contour line.

Under the Lorentzian metric in Eq. (4), the contour lines in Fig. 8(b) where all points are equidistant to the starting point in $(0, 0)$, are not circular. The contour lines are asymptotically equal to the wave fronts in airplane boarding when $N \to \infty$. The path which is orthogonal to the contour lines is the longest curve (geodesic) under the Lorentzian metric, and it will take the same amount of time to reach any point on a contour line, i.e., all passengers on this line (wave) sit down simultaneously.

The starting point of the longest chain corresponds to the first passenger to sit during the boarding process. Likewise, the end point of the longest chain corresponds to the last passenger seated. If the first passenger in the queue is seated at the first row ($q = 0$, $r = 0$), that passenger can block all other passengers and will be in the first wave front. If the last passenger in the queue is seated at the last row ($q = 1$, $r = 1$), that passenger is blocked by all the other passengers and must be in the last wave front. For large N and uniform distribution there will be passengers with (q, r) coordinates arbitrarily close to those points. Hence, the curve that approximates the longest chain when N is large should be the longest continuous path under the Lorentzian metric between $(0, 0)$ and $(1, 1)$ within the (q, r) unit square. Examples of such longest paths in airplane boarding are shown in Figs. 3 and 5 and their expressions are given in Appendix A.

V. ASYMPTOTIC BOARDING TIME

A. Asymptotic boarding time with one group

A cornerstone result in the causal set approach is a limiting result by Myrheim [10] which links the number of elements in the longest chain with the length of the longest curve, up to a scaling factor depending only on the dimension of the domain.

The scaling factor was found independently by Vershik and Kerov [22] and Logan and Shepp [23] for dimension 2, the same dimension as in airplane boarding. The $\sqrt{N}$ law follows from simple subadditivity arguments.

Let τ be the common aisle-clearing time for each of the passengers, N the number of passengers, and $\max_r L(r)$ the length in Eq. (4) of the longest causal curve $r(q)$ between $(0, 0)$ and $(1, 1)$ within the (q, r) unit square. For large N, the longest chain follows closely the trajectory of the longest causal curve (see Fig. 5) and both satisfy the blocking relation in Eq. (2). A generalization of a result of Deuschel and Zeitouni [24] states that the boarding time converges to a multiple of the length of the longest causal curve [7]

$$\frac{T}{\sqrt{N}} \overset{\text{a.s.}}{\to} 2\tau \max_r L(r).$$

From this, the asymptotic average boarding time is given by

$$\langle T \rangle \sim 2\tau \sqrt{N} \max_r L(r) \equiv \hat{T}. \tag{5}$$

The asymptotic boarding time $\hat{T}$ is a leading term and has been shown to overestimate the finite-N average boarding time $\langle T \rangle$ by a relative error of $o(N^{-1/4})$ [4]. Still, the relative ranking of boarding policies has shown to be maintained for small N, as demonstrated by discrete-event simulations in Fig. 4(b) and Ref. [1].

The procedure for computing the length of the longest curves is presented in Sec. VI. For the random boarding examples in Figs. 1 and 5(a), we set $\tau = 1$. In Fig. 5(a) $k = 4$, $N = 4000$, and $r(q) \equiv 0$ for $q < 0.83$ and $r(q) = 4(e^{-2k(1-q)} - e^{-k(1-q)}) + 1$ for $q \geqslant 0.83$. The length of this curve is 2.153, and from Eq. (5), $\hat{T} = 272$. In Fig. 1, $k = 1$ and $N = 8$ and a corresponding curve gives $\hat{T} = 7.4$. The actual boarding times are 232 and 3 for the two examples, respectively. This illustrates that the asymptotic estimate $\hat{T}$ in Eq. (5) can be inaccurate for small N, but improves as N increases.

B. Asymptotic boarding time for two groups with different aisle-clearing times

In this paper we consider policies where all the slow (or fast) passengers are placed in the first part of the queue. Hence, the aisle-clearing time is different for the two groups of passengers. Let p be the fraction of slow passengers. In addition, τ_S and τ_F are the aisle-clearing times for slow and fast groups, respectively.

The asymptotic average boarding time in Eq. (5) must be modified to reflect the fact that the queue now consists of two separate groups with different aisle-clearing times. The aisle-clearing time can be parametrized according to the queue position $\tau = \tau(q)$. For the slow-first policy, $\tau(q) = \tau_S$ for $q \leqslant p$ and $\tau(q) = \tau_F$ for $q > p$, and correspondingly for the fast-first policy.

The aisle-clearing time can be thought of as the proper time (Lorentzian metric length) between two successive passengers (events) in a chain. The boarding time is no longer given by the maximal length of a causal chain, but rather by the causal chain with maximal weight. The definition of length in Eq. (4) must be scaled to reflect that the aisle-clearing time of each

passenger depends on the queue position. Following [7], the curve weight (proper time) of a causal curve is defined by

$$W(r) = \int_{q_0}^{q_1} \tau(q) \sqrt{r'(q) + k[1 - r(q)]}\, dq, \tag{6}$$

where $\tau(q)$ also can be considered a weight function applied to the integrand in Eq. (4). When there are two groups, as in, e.g., the slow-first policy, the curve weight on the interval $q \in (0, 1)$ is given by

$$\begin{aligned} W_{\text{SF}}(r) &= \int_0^p \tau_S \sqrt{r'(q) + k[1 - r(q)]}\, dq \\ &\quad + \int_p^1 \tau_F \sqrt{r'(q) + k[1 - r(q)]}\, dq \\ &= \tau_S L_S(r) + \tau_F L_F(r), \end{aligned} \tag{7}$$

where L_S and L_F are curve lengths as defined in Eq. (4).

The boarding time is (for $N \gg 1$) proportional to the longest (i.e., heaviest) curve $r(q)$ from $(0, 0)$ to $(1, 1)$ within the unit square:

$$\langle T \rangle \sim 2\sqrt{N} \max_r W(r) \equiv \tilde{T}. \tag{8}$$

An additional constraint on the curve is that $r'(q)$ must be continuous whenever $\tau(q)$ is continuous. The result in Eq. (8) is used in Sec. VI to derive analytical expressions for the expected boarding time for both the slow-first and the fast-first boarding policies (again, for $N \gg 1$).

When the aisle-clearing times vary within the groups, τ_S and τ_F in Eq. (7) can be replaced by corresponding parameters that take the variations into account. Although explicit expressions for the parameters are seldomly available, they can be estimated by the square root of the second moment of τ_S and τ_F, respectively [20]. A detailed treatment of this aspect is left for future work.

C. Airplane boarding and geometric optics

The result above for airplane boarding states that the boarding time can be derived from the curve (geodesic) that maximizes the proper time (curve weight) in space-time under an appropriate Lorentzian metric. In geometric optics Fermat's principle states that light will travel between two points along a path that minimizes the travel time, which is a function of the local index of refraction.

In Eqs. (6) and (7) the aisle-clearing time τ plays the same role in the Lorentzian space for airplane boarding as the refractive index in the Euclidean space for light. This can be seen in the qr diagrams in Figs. 3(c), 3(d) and 5(b), where the longest curve is refracted at the boundary between different groups of passengers with different aisle-clearing time. While light moves in straight lines in homogeneous media, the longest curve under the Lorentzian metric has a curved shape in the space-time domains filled with passengers with equal aisle-clearing time. Interestingly, when $k = 0$, the curves become straight also under the Lorentzian metric and refraction occurs at the border between passenger groups according to a principle similar to Snell's law.
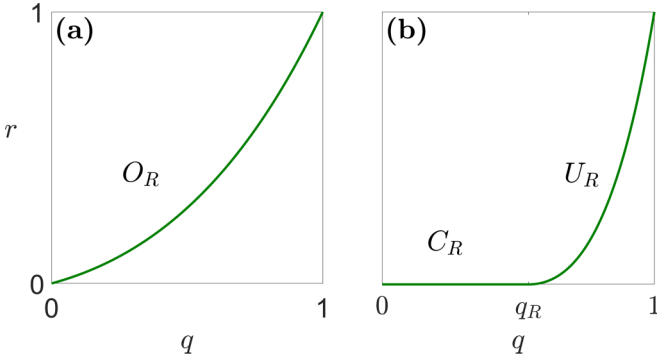
FIG. 9. The shape of the longest curve for random boarding can be either ordinary or piecewise. (a) Ordinary-type curve (O_R) when $0 < k \leqslant \ln(2)$. (b) Piecewise curve (P_R) consisting of a constant function (C_R) and an upward-going ordinary-type curve (U_R) when $k > \ln(2)$.

VI. BOARDING TIMES FOR SLOW-FIRST AND FAST-FIRST POLICIES

We now turn to computing the asymptotic average boarding time in Eq. (8) for $k > 0$, with p and $C \equiv \tau_F/\tau_S$ both in the range $(0, 1)$. We show that the slow-first policy is more efficient than the fast-first policy in the entire (k, p, C) parameter space in the large-N limit ($N \to \infty$). Comparisons to simulation results for smaller N are also made.

A. Analysis of the random boarding policy

To better explain our analytical approach, we first illustrate the computations of the average boarding time of the random boarding policy. With one group, $\tau(q) \equiv \tau$ when all passengers have the same aisle-clearing time. The curve weight in Eq. (6) becomes $W(r) = \tau L(r)$. The curve length $L(r)$ in Eq. (4) can be maximized using the variational method. This leads to general solutions of the form $r^*(q) = ae^{2kq} + be^{kq} + 1$ when $k > 0$.

The constants a and b are determined using the values at the start and end points: $r^*(0) = 0$ and $r^*(1) = 1$. A typical shape is shown in Fig. 9(a) for $k \leqslant \ln(2)$. The resulting maximal length of what we call an ordinary-type curve is, by Eq. (4),

$$L(r^*) = \sqrt{(e^k - 1)/k} \equiv O_R. \tag{9}$$

However, when $k > \ln(2)$, an ordinary-type curve r^* will extend below the (q, r) unit square. Since the curve should be within the unit square,[5] the first part of the curve should be horizontal at value zero along the q axis [C_R in Fig. 9(b)]. The remaining part is an upward-going ordinary-type curve (U_R). Continuity of r and r' in the transition point $q = q_R$ between the C_R and U_R curves and that in the end point $r(1) = 1$ determine the values of a and b in the U_R curve (see Appendix A for details). This gives $q_R = 1 - \ln(2)/k$, and the total length of the resulting piecewise curve r_P^* is [for

$k > \ln(2)$]

$$L(r_P^*) = \frac{1}{\sqrt{k}}[k - \ln(2) + 1] \equiv P_R. \tag{10}$$

Hence, the expected boarding time with random boarding and equal aisle-clearing time τ is by leading order given by Eq. (5), which gives

$$\hat{T}_R = \begin{cases} 2\tau\sqrt{\frac{N}{k}}\sqrt{e^k - 1}, & 0 < k \leqslant \ln(2) \\ 2\tau\sqrt{\frac{N}{k}}[k - \ln(2) + 1], & \ln(2) < k. \end{cases} \tag{11}$$

B. Analysis of the slow-first policy

The curve weight for the slow-first policy is given by Eq. (7). The longest curve must be continuous, but does not have to be smooth at the crossing point $(p, r(p))$ between the regions of slow and fast passengers in the qr diagram. If we fix the crossing height $r(p) = \delta$, the longest curves in each part of the qr diagram must be either ordinary type or piecewise, as for the single group policy in Sec. VI A.

The length L_S of the longest curve in the first part of the diagrams in Fig. 10 is the length of a piecewise curve ($L_S = P_S$) if $\delta < \delta_S \equiv (e^{kp} - 1)^2$ and of an ordinary-type curve ($L_S = O_S$) if $\delta \geqslant \delta_S$. A similar parameter $\delta_F = \max\{0, 1 - 2e^{-k(1-p)}\}^2$ determines the type of longest curve in the second part of the qr diagrams. Explicit expressions for the curves (O_S, P_S, etc.) are given in Appendix A.[6]

Depending on the value of the fixed crossing height δ [relative to the values of $\delta_S(k, p)$ and $\delta_F(k, p)$], the resulting total curves can be one of four different combinations of each of these curves. Let $W_{SF}(\delta)$ be the weight of the longest piecewise curve for a fixed crossing height δ (the dependence on k, p, τ_S, and τ_F is suppressed in the following):

$W_{SF}(\delta; k, p, \tau_S, \tau_F)$

$$= \begin{cases} W_{SF1}(\delta) = \tau_S P_S(\delta) + \tau_F P_F(\delta), & \delta < \min\{\delta_S, \delta_F\} \\ W_{SF2}(\delta) = \tau_S P_S(\delta) + \tau_F O_F(\delta), & \delta_F \leqslant \delta < \delta_S \\ W_{SF3}(\delta) = \tau_S O_S(\delta) + \tau_F P_F(\delta), & \delta_S \leqslant \delta < \delta_F \\ W_{SF4}(\delta) = \tau_S O_S(\delta) + \tau_F O_F(\delta), & \max\{\delta_S, \delta_F\} \leqslant \delta. \end{cases} \tag{12}$$

To find the longest curve, we must compute the $\delta = \delta^*$ that maximizes $W_{SF}(\delta)$. The function $W_{SF}(\delta)$ is differentiable and has negative curvature when $\delta \in (0, 1)$. This means that there is never more than one local maximum in the domain.

Moreover, each of the four subfunctions in Eq. (12) has a maximum point $\delta_i^* \in [0, 1]$, $i \in \{1, 2, 3, 4\}$, with maximum value W_{SFi}^*. Each of these are global maximum points for $W_{SF}(\delta)$ if and only if δ_i^* lies within the respective subdomain in Eq. (12). Hence, the weight of the longest curve for the slow-first policy is, e.g., given by $W_{SF}^* = W_{SF1}^*$ when $\delta_1^* < \min\{\delta_S, \delta_F\}$.

---

[5]All passengers in the longest chain are within the unit square.

[6]The same curves are used for the fast-first policy, only by exchanging $p \leftrightarrow 1 - p$ and $\tau_S \leftrightarrow \tau_F$.
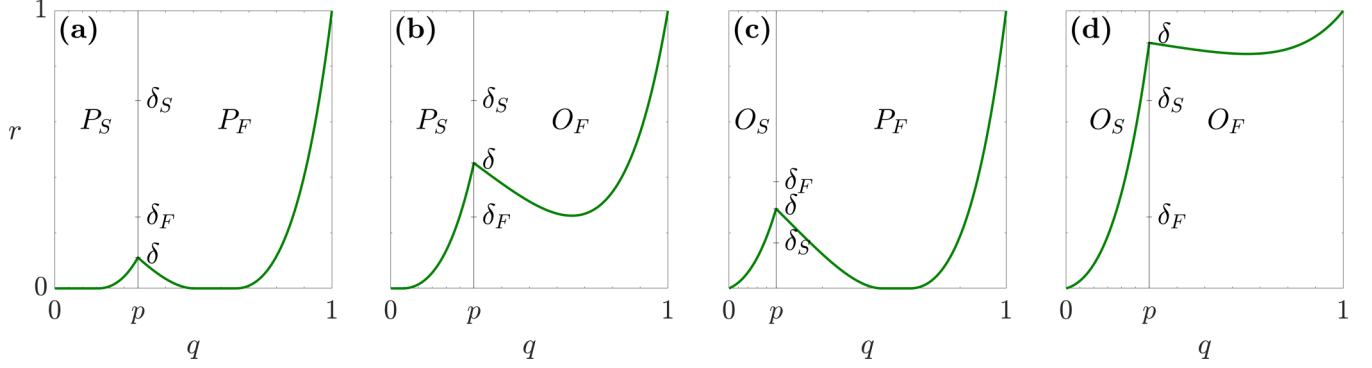
FIG. 10. For fixed crossing height δ, the shape of the longest curve for slow-first can be either piecewise (P_S, P_F) or ordinary (O_S, O_F) in both the slow and the fast regions, respectively. Hence, the total curve can take four different shape types. The subfunction that defines the weight $W_{\mathrm{SF}}(\delta)$ in Eq. (12) depends on the values of $\delta_S(k, p)$ and $\delta_F(k, p)$ (expressions are given in the text): (a) W_{SF1}, $\delta < \min\{\delta_S, \delta_F\}$; (b) W_{SF2}, $\delta_F \leqslant \delta < \delta_S$; (c) W_{SF3}, $\delta_S \leqslant \delta < \delta_F$; and (d) W_{SF4}, $\max\{\delta_S, \delta_F\} \leqslant \delta$. The δ that maximizes $W_{\mathrm{SF}}(\delta)$ also depends on the relative aisle-clearing time $C = \tau_F/\tau_S$.

The maximum points $\delta = \delta_i^*$ for each of the subfunctions in Eq. (12) yield the following global maxima W_{SF}^* of W_{SF}:

$$W_{\mathrm{SF1}}^* = \frac{\tau_S}{\sqrt{k}} \left[kp(1 - C) + kC + 1 + C \ln\left(\frac{C}{1 + C}\right) \right.$$
$$\left. - \ln\left(\frac{2}{1 + C}\right) \right], \quad \max\{C_2, C_1\} \leqslant C$$

$$W_{\mathrm{SF2}}^* = \frac{\tau_S}{\sqrt{k}} \left[kp + 1 - \ln\left(\frac{2}{1 + C^2(e^{k(1-p)} - 1)}\right) \right],$$
$$C_3^2 \leqslant C^2 \leqslant C_1^2$$

$$W_{\mathrm{SF3}}^* = \frac{\tau_S}{\sqrt{k}} \left(\sqrt{(1 - e^{-kp})(e^{kp} - 1 + \delta_3^*)} \right.$$
$$\left. + C\{k(1 - p) + 1 + \sqrt{\delta_3^*} + \ln[(1 - \sqrt{\delta_3^*})/2]\} \right),$$
$$C_4^2 \leqslant C^2 \leqslant C_2^2$$

$$W_{\mathrm{SF4}}^* = \frac{\tau_S}{\sqrt{k}} \sqrt{(e^{kp} - 1) + C^2(e^k - e^{kp})},$$
$$C^2 \leqslant \min\left\{C_3^2, C_4^2\right\}. \tag{13}$$

Here $C \equiv \tau_F/\tau_S \in (0, 1)$. Since $\delta_i^*(k, p, C)$, $\delta_S(k, p)$, and $\delta_F(k, p)$ are functions of (k, p, C), the conditions on $\delta = \delta_i^*$ in Eq. (12) have been rewritten as conditions on C, where $C_1 \equiv (e^{k(1-p)} - 1)^{-1}$, $C_2 \equiv 2e^{-kp} - 1$, $C_3^2 \equiv (2 - e^{kp})/(e^k - e^{kp})$, and $C_4^2 \equiv 4(e^{kp} - 1)/[e^{2k} - 4(e^k - e^{kp})]$.[7]

The subdomains where the conditions in Eq. (13) are satisfied are shown in the top row of Fig. 11 for $k \in \{0.5, 1, 4\}$. For example, for $kp > \ln(2)$, the conditions are simplified such that W_{SF1}^* is the maximum when $C \geqslant C_1$ and W_{SF2}^* when $C \leqslant C_1$.

The maximal weight W_{SF}^* in Eq. (13) is used to calculate the corresponding asymptotic boarding time $\bar{T}_{\mathrm{SF}}$ in Eq. (8). In Fig. 12(a) comparisons of the asymptotic boarding time for the slow-first policy with simulation results for $N \leqslant 240$ show that the asymptotic result in Eq. (8) tends to overestimate

---

[7]The expression for the maximum point δ_3^* for $W_{\mathrm{SF3}}(\delta)$ is given in Appendix A.

the boarding time, but the relative ranking between different parameter settings is maintained.

C. Analysis of the fast-first policy

The same procedure as in the preceding section can be repeated for the fast-first policy by exchanging $p \leftrightarrow 1 - p$, $\tau_S \leftrightarrow \tau_F$, and $C \to 1/C$. For $W_{\mathrm{FF1}}(\delta)$ (which corresponds to W_{SF1}), $W_{\mathrm{FF1}}'(\delta) < 0 \, \forall \, \delta \in (0, 1)$. Consequently, the weight of the longest curve is given for $\delta = 0$ and

$$W_{\mathrm{FF}}^* = W_{\mathrm{FF1}}(0) = \frac{\tau_S}{\sqrt{k}} [kp(1 - C) + kC + 1 - \ln(2)]. \tag{14}$$

The weight $W_{\mathrm{FF}}(\delta)$ has negative curvature and $\delta = 0$ is therefore the global maximum point for W_{FF} as long as $\delta = 0$ satisfies the conditions corresponding to the subdomains for FF1 in Eq. (12). For fast first, $\delta_S(k, 1 - p) = (1 - e^{k(1-p)})^2 > 0$. Moreover, $\delta_F(k, 1 - p) = \max\{0, 1 - 2e^{-kp}\}^2$ is also positive when $kp > \ln(2)$. This means that when $kp > \ln(2)$, the condition $\delta = 0 < \min\{\delta_S(k, 1 - p), \delta_F(k, 1 - p)\}$ is satisfied, and hence $W_{\mathrm{FF}}^* = W_{\mathrm{FF1}}^*$. The corresponding subdomain is indicated in the bottom row of Fig. 11.

For $kp \leqslant \ln(2)$, results for the fast-first policy corresponding to those in Eq. (13) are given in Appendix A. In Fig. 12(b) comparisons of the asymptotic boarding time for the fast-first policy with simulation results for $N \leqslant 240$ are similar to those for the slow-first policy.

D. Comparing slow-first and fast-first policies

The slow-first policy outperforms the fast-first policy for all values of $k > 0$ and $p, C \in (0, 1)$. Here explicit results are only shown for $kp > \ln(2)$. The derivation of the other results are left for Appendices A and B, and the results are summarized in Sec. III with comparisons to finite-N discrete-event simulations.

Based on the results in Eqs. (13) and (14), $W_{\mathrm{FF}}^* - W_{\mathrm{SF}}^*$ is, for $kp > \ln(2)$, given by

$$W_{\mathrm{FF}}^* - W_{\mathrm{SF}}^* = \frac{\tau_S}{\sqrt{k}} \left[\ln\left(\frac{1}{1 + C}\right) - C \ln\left(\frac{C}{1 + C}\right) \right],$$
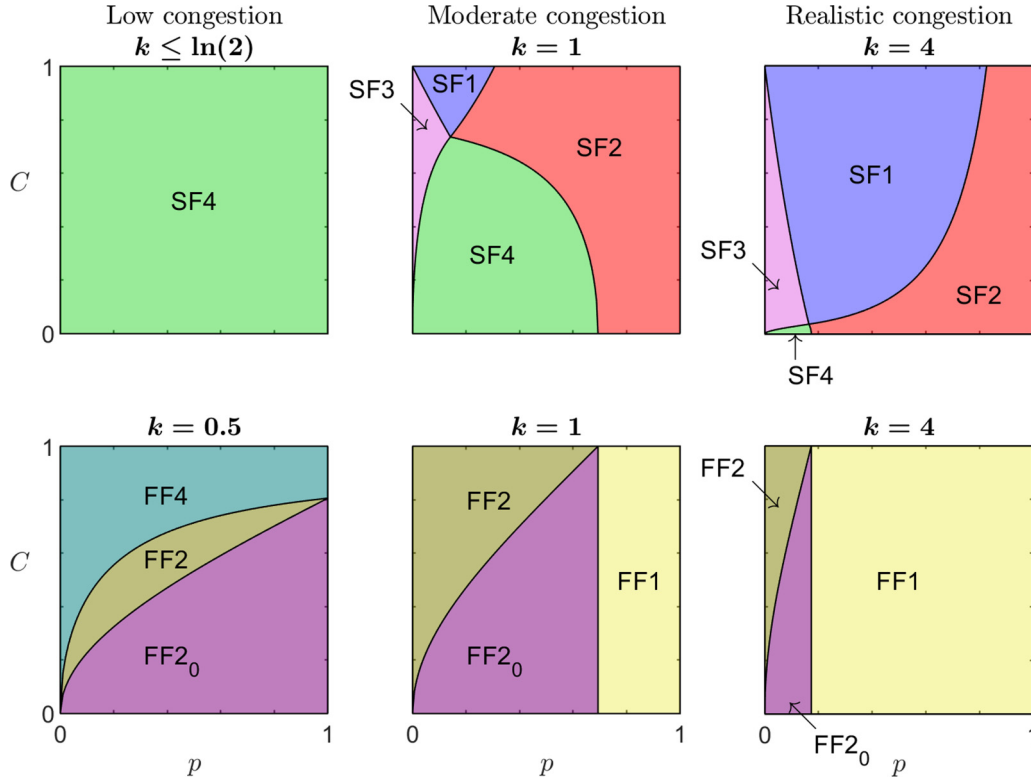$$C_1 \leqslant C \tag{15}$$

FIG. 11. The top row shows the subdomains of the (p, C) unit square where the slow-first boarding time is represented by the different subfunctions in Eq. (13). The bottom row shows the corresponding subdomains for the fast-first policy in Eq. (A3).

$$W_{\mathrm{FF1}}^{*} - W_{\mathrm{SF2}}^{*} = \frac{\tau_S}{\sqrt{k}} \{Ck(1-p) - \ln[1 + C^2(e^{k(1-p)} - 1)]\},$$

$$C \leqslant C_1, \qquad (16)$$

where $C_1 \equiv (e^{k(1-p)} - 1)^{-1}$. In Appendix B, both of these are shown to be positive, meaning that the fast-first policy yields a longer average boarding time than the slow-first policy.

The difference approaches zero for both $C \to 0$ and $C \to 1$. That $C \to 0$ means that the fast passengers are much faster than the slow ones. Then it does not matter who enters first since the fast fraction of the queue will sit down immediately anyway. When $C \to 1$, both groups have the same speed and the two policies should not differ.

The difference also vanishes when $k \to \infty$. This means that each passenger takes up so much room that the first
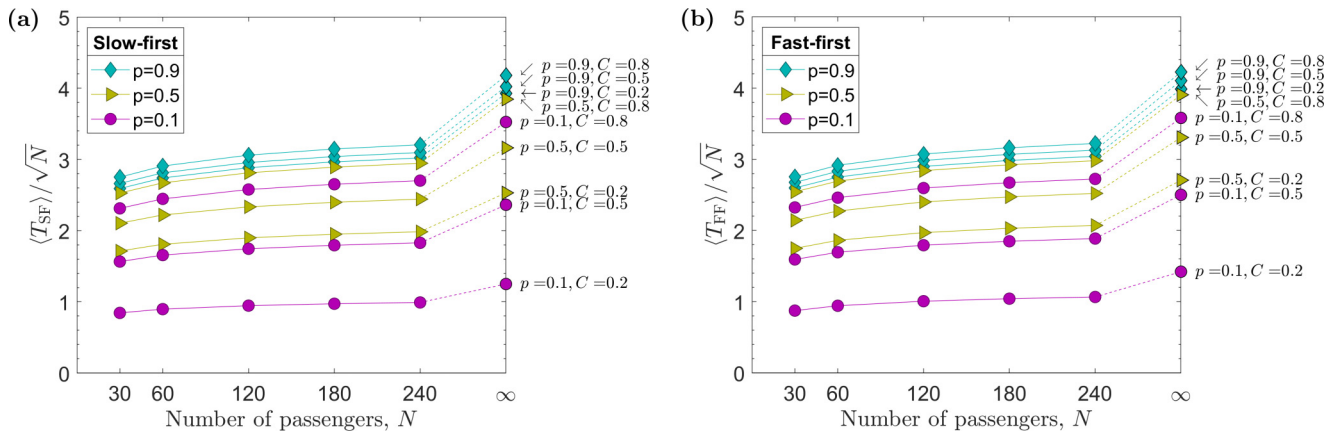

FIG. 12. Average boarding time estimates for the (a) slow-first and (b) fast-first policies for different (k, p, C) parameter settings. Simulation results for an increasing number of passengers are compared to the asymptotic results for all combinations of parameter values $p \in \{0.1, 0.5, 0.9\}$ and $C \in \{0.2, 0.5, 0.8\}$. Here $k = 4$ and the accuracy is ± 0.002 (as a result of 10^6 runs for each finite-N data point). The rightmost points are asymptotic values.

group will be seated before the next group enters the airplane. Consequently, the sequencing of the groups does not matter.

When $1 - p \to 0$ the difference also approaches zero. This means that when the fraction $1 - p$ of fast passengers vanishes, there is no difference between slow first and fast first. This is expected since the queue then reduces to slow passengers only.

VII. DISCUSSION AND OUTLOOK

In this paper, airplane boarding has been analyzed using Lorentzian geometry, which is exact in the limit of infinitely many passengers. We have shown that the boarding time with the slow-first policy on average is lower than for the fast-first policy in the large-N limit. The reason for slow-first superiority is that this policy better utilizes parallelism, extending the time window during which the first fast passengers can sit down simultaneously with the last slow passengers. With the fast-first policy, on the other hand, the last fast passengers rapidly take their seats, with less time remaining for simultaneous seating of the slow passengers.

Simulations we performed confirm that our asymptotic result still applies for a lower number of passengers N of the order of hundreds, as in medium-sized airplanes. There are also some basic similarities with back-to-front boarding, which is a beneficial policy when the congestion factor k is small, since then the two groups can sit down in parallel. In general, the slow-first policy has better parallelism compared to fast first, but of a different kind which is much less congestion dependent than the back-to-front policy.

A future extension of this work can include increasing the variability to more than just two groups of passengers' aisle-clearing time. For an inhomogeneous group of passengers with different aisle-clearing times, one can attach an effective aisle-clearing time which measures by which factor the group as a total slows the boarding process [20]. This parameter plays the role of an effective refractive index, where each group could consist of both slow and fast passengers. Preliminary results indicate that the slow-first policy is always better than the random boarding policy, in which the fast and slow passengers are randomly distributed in one group. Assuming this is the case, one can argue that both the slow and the fast group should be divided over again according to speed. Eventually, the slowest passenger would be in front of the queue and the fastest passenger in the back of the queue. Whether this would reduce the boarding time is still an open question.

For airlines it is not only the average boarding time that matters. The frequency of departure delays due to an unexpectedly long boarding time could also be of relevance. Then the percentage of boardings exceeding a certain duration would be an appropriate measure. The same types of results as we presented here must then be obtained in terms of percentiles. Such percentiles can be estimated using the Tracy-Widom distribution [7], but requires extensive calculations.

The boarding process at airports is certainly more complicated than depicted in our models. However, recent empirical research shows that such models explain real data rather well [25]. The methods used to compute the asymptotic boarding time can also be used when assumptions and policies deviate from those used in this paper. The modifications needed to estimate boarding time when there is, e.g., a half-empty airplane or varying passenger widths w are described in the Supplemental Material of Ref. [21]. Moreover, in this paper we do not take into account that a window seat passenger takes a longer time to settle down if the middle-seat or aisle-seat passenger in the same row is already seated. How this affects the boarding time can also be taken into account within our framework, but the computational details are more complex. However, this would let us compare the random boarding policy with the group-based window-middle-aisle boarding policy which is applied by some airlines.

Will the fast passengers accept the slow-first policy? Some passengers prefer to spend as little time as possible in confined spaces and are very happy to be the last passenger to enter the airplane. However, other might dislike waiting or the idea that the "troublesome" passengers with much carry-on luggage are occupying the overhead lockers, leaving little room for the coats and smaller items of the light-traveling passengers. This could potentially lead to an unintended consequence where passenger would have an incentive to bring more carry-on luggage on board, leading to a kind of tragedy of the commons scenario where everybody will wait longer.

ACKNOWLEDGMENTS

APPENDIX A: CALCULATION DETAILS

1. One group: Random boarding

In this section we show more details of the computations behind the maximized curve lengths of the random boarding policy in Eqs. (9) and (10). The problem is broken down by first noting that the general solution of maximized curve length (4) between any two points is of the form $r^*(q) = ae^{2kq} + be^{kq} + 1$. There is one such function (which we refer to by its length O_R) that goes between $(0, 0)$ and $(1, 1)$. When $k > \ln(2)$ this curve would dip below the q axis and only the last part U_R of a piecewise smooth function would be of the mentioned form. The parameters a and b for these curves (shown in Fig. 9) are given by

$$a = (e^k - 1)^{-1}, \quad b = -ae^k, \quad q \in (0, 1),$$
$$k \leqslant \ln(2) \quad \text{for the } O_R \text{ curve,}$$
$$a = 4e^{-2k}, \quad b = -4e^{-k}, \quad q \in (1 - \ln(2)/k, 1),$$
$$k > \ln(2) \quad \text{for the } U_R \text{ curve.}$$

In order to compute the length of each of the curves in Fig. 9, we use that the length $L(r)$ in Eq. (4) of a function $r(q)$ between two points q_1 and q_2 is given by

$$L(r) = \sqrt{\frac{a}{k}}(e^{kq_2} - e^{kq_1}) \quad \text{when } r(q) = ae^{2kq} + be^{kq} + 1,$$

$$L(r) = \sqrt{k}(q_2 - q_1) \quad \text{when } r(q) \equiv 0.$$

(A1)

This gives the length of both r^* (O_R) and r_P^* ($P_R = C_R + U_R$) in Eqs. (9) and (10), respectively.

2. Two groups: Slow-first policy

In this section we show more details of the computations behind the maximized curve weight of the slow-first policy in Eq. (7) (and the corresponding one for the fast-first policy). The problem is broken down by first noting that the general solution of maximized curve length (4) between any two points is of the form $r^*(q) = ae^{2kq} + be^{kq} + 1$. There is one such function in each of the two regions shown in Fig. 10(d) when the crossing height $r^*(p) = \delta$ is fixed. The O_S curve goes between the points $(0, 0)$ and (p, δ), while the O_F curve continues to $(1, 1)$. The parameters a and b are given by

$$a = \frac{e^{kp} - 1 + \delta}{e^{2kp} - e^{kp}}, \quad b = -(a+1), \quad q \in (0, p),$$

$$\delta \geqslant \delta_S \quad \text{for the } O_S \text{ curve,}$$

$$a = \frac{1 - \delta}{e^{kp}(e^k - e^{kp})}, \quad b = -ae^k, \quad q \in (p, 1),$$

$$\delta \geqslant \delta_F \quad \text{for the } O_F \text{ curve.}$$

The restriction on δ is in order to avoid that the curve dips below the q axis. For example, when $\delta < \delta_S$, the O_S curve in the first region in Figs. 10(c) and 10(d) would have dipped below the q axis. To maintain a positive continuous smooth curve as in Figs. 10(a) and 10(b), a part of the curve must be constant along the q axis. The remaining parts of these piecewise curves are ordinary-type curves that are connected smoothly to the constant part of the curve. They have the parameters

$$a = e^{-2kp}(1 + \sqrt{\delta})^2, \quad b = -2e^{-kq_S},$$

$$q \in (q_S, p) \quad \text{(last part of the } P_S \text{ curve),}$$

$$a = e^{-2kp}(1 - \sqrt{\delta})^2, \quad b = -2e^{-kq_F},$$

$$q \in (p, q_F) \quad \text{(first part of the } P_F \text{ curve),}$$

$$a = 4e^{-2k}, \quad b = -4e^{-k},$$

$$q \in (1 - \ln(2)/k, 1) \quad \text{(last part of the } P_F \text{ curve),}$$

where $q_S = p - \ln(1 + \sqrt{\delta})/k$ and $q_F = p - \ln(1 - \sqrt{\delta})/k$.

The length of each of the curves in Fig. 10 is found by Eq. (A1). This gives

$$O_S(\delta) = \frac{1}{\sqrt{k}}\sqrt{(e^{-kp} - 1)(e^{kp} - 1 + \delta)},$$

$$O_F(\delta) = \frac{1}{\sqrt{k}}\sqrt{(1 - \delta)(e^{k(1-p)} - 1)},$$

$$P_S(\delta) = \frac{1}{\sqrt{k}}[kp + \sqrt{\delta} - \ln(1 + \sqrt{\delta})],$$

$$P_F(\delta) = \frac{1}{\sqrt{k}}[k(1 - p) + (1 + \sqrt{\delta}) + \ln(1 - \sqrt{\delta}) - \ln(2)].$$

For slow first, these are combined to $W_{\mathrm{SF}}(\delta)$ according to Eq. (12). Since $W_{\mathrm{SF}}(\delta)$ is differentiable and has negative

curvature when $\delta \in (0, 1)$ there is never more than one local maximum on the domain.

To show the differentiability, we need continuity of $W_{\mathrm{SF}}'(\delta)$ in the transition points when $\delta = \delta_S, \delta_F$. For $\delta_S < \delta_F$ we need that $W_{\mathrm{SF1}}'(\delta_S) = W_{\mathrm{SF3}}'(\delta_S)$ and $W_{\mathrm{SF3}}'(\delta_F) = W_{\mathrm{SF4}}'(\delta_F)$. For $\delta_F < \delta_S$, we must require that $W_{\mathrm{SF1}}'(\delta_F) = W_{\mathrm{SF2}}'(\delta_F)$ and $W_{\mathrm{SF2}}'(\delta_S) = W_{\mathrm{SF4}}'(\delta_S)$. This reduces to showing that $P_S'(\delta_S) = O_S'(\delta_S)$ and $P_F'(\delta_F) = O_F'(\delta_F)$, which is straightforward. The negative curvature of W_{SF} follows from $W_{\mathrm{SF}i}''(\delta) < 0 \,\forall\, \delta \in (0, 1)$ for $i = 1, 2, 3, 4$.

The δ's that are maximizing each of the subfunctions in Eq. (12) are given by

$$\sqrt{\delta_1^*} = \frac{1 - C}{1 + C},$$

$$\sqrt{\delta_2^*} = \frac{1 - C^2(e^{k(1-p)} - 1)}{1 + C^2(e^{k(1-p)} - 1)},$$

$$\sqrt{\delta_3^*} = \frac{\sqrt{(e^{kp} - 1)(1 - C^2)} - C(e^{kp} - 1)}{\sqrt{(e^{kp} - 1)(1 - C^2)} + C}, \tag{A2}$$

$$\delta_4^* = \frac{1 - C^2(e^k - e^{kp})}{1 + C^2(e^k - e^{kp})/(e^{kp} - 1)}.$$

If $\delta_i^* = \arg\max_{\delta \in (0,1)} W_{\mathrm{SF}i}(\delta)$ lies in the subdomain where $W_{\mathrm{SF}} = W_{\mathrm{SF}i}$, then δ_i^* is the global maximum of W_{SF} due to the negative curvature of $W_{\mathrm{SF}}(\delta)$. Inserting these into Eq. (12) gives W_{SF}^* in Eq. (13).

When $k > \ln 2$, the boundaries shown in the top row of Fig. 11 meet at the vortex point (p^*, C^*), where $C^* = 2e^{-k}$ and $p^* = \frac{1}{k}\ln(\frac{2}{1+C^*})$. The boundaries are given by

$$C_1 = \frac{1}{e^{k(1-p)} - 1}, \quad p \in [p^*, 1 - (\ln 2/k)] \quad \text{(SF1-SF2)},$$

$$C_2 = 2e^{-kp} - 1, \quad p \in [0, p^*] \quad \text{(SF1-SF3)},$$

$$C_3 = \sqrt{\frac{2 - e^{kp}}{e^k - e^{kp}}}, \quad p \in [p^*, \ln 2/k] \quad \text{(SF2-SF4)},$$

$$C_4 = 2\sqrt{\frac{e^{kp} - 1}{e^{2k} - 4(e^k - e^{kp})}}, \quad p \in [0, p^*] \quad \text{(SF3-SF4)}.$$

3. Two groups: Fast-first policy

The same procedure as in the preceding section can be repeated for the fast-first policy by exchanging $p \leftrightarrow 1 - p$, $\tau_S \leftrightarrow \tau_F$, and $C \to 1/C$. The δ's maximizing each of the subfunctions of $W_{\mathrm{FF}}(\delta) \equiv W_{\mathrm{SF}}(\delta; k, 1 - p, \tau_F, \tau_S)$ in Eq. (12) are given by

$$\sqrt{\tilde{\delta}_1^*} = 0,$$

$$\sqrt{\tilde{\delta}_2^*} = \begin{cases} 0, & C \leqslant \tilde{C}_1 \\ \frac{C^2 - (e^{kp} - 1)}{C^2 + (e^{kp} - 1)}, & C > \tilde{C}_1, \end{cases}$$

$$\tilde{\delta}_4^* = \frac{C^2 - (e^k - e^{k(1-p)})}{C^2 + (e^k - e^{k(1-p)})/(e^{k(1-p)} - 1)},$$

where $\tilde{C}_1 = \sqrt{e^{kp} - 1}$. Here $\tilde{\delta}_3^*$ is not computed since this will never be a global maximum for $W_{FF}(\delta)$. Inserting $\tilde{\delta}_1^*$ into $W_{FF}(\delta)$ gives $W_{FF}^* = W_{FF1}^*$ in Eq. (14) when $kp > \ln(2)$. When $kp \leqslant \ln(2)$, $\tilde{\delta}_2^*$ and $\tilde{\delta}_4^*$ can be inserted into $W_{FF}(\delta)$, which gives the expression for W_{FF}^*,

$$W_{FF2_0}^* = \frac{\tau_S}{\sqrt{k}}[Ck(1-p) + \sqrt{e^{kp} - 1}], \quad C^2 \leqslant \tilde{C}_1^2,$$

$$W_{FF2}^* = \frac{\tau_S}{\sqrt{k}}\{Ck(1-p) + C[1 - \ln(2)]$$

$$+ C\ln[1 + (e^{kp} - 1)/C^2]\}, \quad \tilde{C}_1^2 < C^2 < \tilde{C}_2^2,$$

$$W_{FF4}^* = \frac{\tau_S}{\sqrt{k}}\sqrt{C^2(e^{k(1-p)} - 1) + (e^k - e^{k(1-p)})}, \quad \tilde{C}_2^2 \leqslant C^2,$$

$$\text{(A3)}$$

where $\tilde{C}_1^2 \equiv e^{kp} - 1$ and $\tilde{C}_2^2 \equiv (e^{kp} - 1)/(2e^{-k(1-p)} - 1)$.

APPENDIX B: PROOFS

There are 13 different combinations of FF and SF subfunctions which we treat in the following sections in order to show that $W_{FF}^* - W_{SF}^* > 0$ [for $k \leqslant \ln(2)$ there are three combinations, for $k > \ln(2) \cap kp > \ln(2)$ there are two combinations, and for $k > \ln(2) \cap kp \leqslant \ln(2)$ there are eight combinations].

1. Proof that $W_{FF}^* - W_{SF}^* > 0$ when $k \leqslant \ln(2)$

Assume $0 < k \leqslant \ln(2)$ and $p, C \in (0, 1)$. The (p, C) subdomains for the subfunctions of W_{SF}^* and W_{FF}^* are illustrated in the diagrams in the left column in Fig. 11. From Eq. (13) $W_{SF}^* = W_{SF4}^* = \frac{\tau_S}{\sqrt{k}}\sqrt{(e^{kp} - 1) + C^2(e^k - e^{kp})}$, and W_{FF}^* is given by Eq. (A3) since W_{FF1}^* cannot be a solution when $k \leqslant \ln(2)$. We now show that $W_{FF}^* - W_{SF}^* > 0$ for all three cases.

Case of FF2$_0$ vs SF4. Set $x \equiv k(1-p) \in (0, \ln(2))$. From Eq. (A3), $C^2 \leqslant e^{kp} - 1$. Then

$$\frac{k}{\tau_S^2}\left(W_{FF2_0}^{*2} - W_{SF4}^{*2}\right) = C^2 k^2(1-p)^2 + 2kC(1-p)\sqrt{e^{kp} - 1}$$

$$- C^2(e^k - e^{kp})$$

$$\geqslant C^2[k^2(1-p)^2 + 2k(1-p)$$

$$- e^k(1 - e^{-k(1-p)})]$$

$$= C^2[x^2 + 2x - e^k(1 - e^{-x})]$$

$$\geqslant C^2[x^2 + 2x - 2(1 - e^{-x})] \equiv C^2 g(x),$$

where we use that $C^2 \leqslant e^{kp} - 1$ and $k \leqslant \ln(2)$ in the first and the second inequality, respectively. In addition, $g(0) = 0$, and since $x \in (0, \ln(2))$, $g'(x) = 2x + 2 - 2e^{-x} > 2x > 0$. Hence $g(x) > 0$ for $x > 0$. Q.E.D.

Case of FF2 vs SF4. Set $x \equiv (e^{kp} - 1)/C^2$ and $z \equiv 2e^{-k(1-p)} - 1$. Since $k \leqslant \ln(2)$, $z \in (0, 1)$. In Eq. (A3) the lower bound $C^2 > e^{kp} - 1$ gives $x \in (0, 1)$. The upper bound $C^2 < (e^{kp} - 1)/(2e^{-k(1-p)} - 1)$ gives $x > z$. Then, since

$k \leqslant \ln(2)$,

$$\frac{\sqrt{k}}{C\tau_S}(W_{FF2}^* - W_{SF4}^*) = k(1-p) + 1 - \ln(2)$$

$$+ \ln[1 + (e^{kp} - 1)/C^2]$$

$$- \sqrt{(e^{kp} - 1)/C^2 + e^k(1 - e^{-k(1-p)})}$$

$$= \ln(2) - \ln(z+1) + 1 - \ln(2)$$

$$+ \ln(1 + x) - \sqrt{x + e^k(1 - z)/2}$$

$$\geqslant 1 + \ln(1 + x) - \ln(1 + z)$$

$$- \sqrt{1 + x - z} \equiv g(x, z).$$

For $x = z$, $g(x, z) = 0$. For any $z \in (0, 1)$, $g(x, z)$ is increasing in x since

$$\frac{\partial g}{\partial x} = \frac{1}{1+x} - \frac{1}{2\sqrt{1+x-z}} > \frac{1}{1+x} - \frac{1}{2} > 0.$$

The first inequality is from $x > z$ and the second from $x < 1$. This means that $g(x, z) > 0$ on the triangle domain $z \in (0, 1)$ and $x \in (z, 1)$. Q.E.D.

Case of FF4 vs SF4. It follows straightforwardly that

$$\frac{k}{\tau_S^2}\left(W_{FF4}^{*2} - W_{SF4}^{*2}\right) = (1 - C^2)(e^{kp} - 1)(e^{k(1-p)} - 1) > 0.$$

2. Proof that $W_{FF}^* - W_{SF}^* > 0$ when $k > \ln(2) \cap kp > \ln(2)$

Assume $kp > \ln(2)$ and $p, C \in (0, 1)$. Then $W_{SF}^* - W_{FF}^*$ is given by the two combinations in Eqs. (15) and (16).

Case of FF1 vs SF1. Set $z \equiv \frac{\tau_S}{\tau_S + \tau_F} \in (\frac{1}{2}, 1)$. We can rewrite Eq. (15) as

$$\frac{\sqrt{k}}{\tau_S + \tau_F}(W_{FF1}^* - W_{SF1}^*)$$

$$= \frac{\tau_S}{\tau_S + \tau_F}\ln\left(\frac{\tau_S}{\tau_S + \tau_F}\right) - \frac{\tau_F}{\tau_S + \tau_F}\ln\left(\frac{\tau_F}{\tau_S + \tau_F}\right)$$

$$= z \ln z - (1 - z)\ln(1 - z) \equiv f(z). \quad \text{(B1)}$$

The function $f(z)$ vanishes at the ends of the interval, i.e., $f(\frac{1}{2}) = \lim_{z \to 1^-} f(z) = 0$. The extremum points are calculated from $f'(z) = 0$ or

$$z(1 - z) = e^{-2}. \quad \text{(B2)}$$

Hence, there is a single extremum point on the interval, located at $z = z^* = \frac{1}{2} + \sqrt{\frac{1}{4} - e^{-2}}$. Moreover, from Eqs. (B1) and (B2) we obtain $f''(z^*) = e^2(1 - 2z^*) < 0$, so this extremum is a maximum. According to these properties, $f(z) > 0$ holds within $\frac{1}{2} < z < 1$.

Case of FF1 vs SF2. Set $x \equiv k(1-p) \in (0, \infty)$. The condition $C \leqslant C_1 \equiv (e^{k(1-p)} - 1)^{-1}$ gives that $x \leqslant \ln(1 + C^{-1}) \equiv x_u$. This means that $x \in (0, x_u]$. Then Eq. (16) can be written as

$$\frac{\sqrt{k}}{\tau_S}(W_{FF1}^* - W_{SF2}^*) = Cx - \ln[1 + C^2(e^x - 1)] \equiv g(x, C).$$

It follows that $g(x, C)$ is positive on the interval $x \in (0, x_u]$, since for any fixed $C \in (0, 1)$, $g(0, C) = 0$, and $g(x, C)$ is

monotonically increasing in x. The latter is due to

$$\frac{\partial g}{\partial x} = C - \frac{C^2 e^x}{C^2(e^x - 1) + 1} = C^2(1 - C)\frac{(1 + C^{-1}) - e^x}{C^2(e^x - 1) + 1}$$

being positive for $x \in (0, x_u]$, except in the upper end when $\frac{\partial g}{\partial x}(x_u, C) = 0$.

3. Proof that $W_{\mathrm{FF}}^* - W_{\mathrm{SF}}^* > 0$ when $k > \ln(2) \cap kp \leqslant \ln(2)$

This is when FF2 and $\mathrm{FF2}_0$ in Fig. 11 overlap with the corresponding parts of the SF diagrams in the middle and right columns of Fig. 11. This gives eight different combinations.

We will now show that the difference $W_{\mathrm{FF}}^* - W_{\mathrm{SF}}^* > 0$ within each of these regions, mainly by considering the partial derivatives with respect to p. We use that the difference is smooth on the boundaries between the subdomains. This can be demonstrated quite straightforwardly by differentiation. Moreover, the difference $W_{\mathrm{FF}}^* - W_{\mathrm{SF}}^* = 0$ on the boundaries of the parameter region $p \times C \in [0, 1] \times [0, 1]$.

The outline of the proof is as follows. First, the difference is positive in the SF3 region, including on the SF3-SF1 and SF3-SF4 boundaries as seen in Fig. 11. Then we show that the difference increases in p on the whole SF1 domain, starting out positively on the SF1-SF3 boundary. Now the difference is positive at the SF3-SF4 border and SF1-SF2 border and it is also positive when p is at its maximum at $p = \ln(2)/k$ (at the FF1 border; see Appendix B 2). We show that the difference has negative curvature in p in both the SF2 and SF4 regions. This gives that the difference must be positive for all values of p on p intervals starting at the SF3-SF4 or SF1-SF2 border and ending at $p = \ln(2)/k$ (using that the difference is smooth between regions).

We first show that the difference is increasing in p in the SF1 region. Then the negative curvature in the SF2 and SF4 regions is shown. To show that the difference is positive in the SF3 domain requires a number of subtleties and is left for Appendix B 3 d.

a. The SF1 subdomain

We show that the differences are increasing when p increases for all values of k, p, and C on the SF1 domain.

Case of $\mathrm{FF2}_0$ vs SF1. From Eqs. (A3) and (13) we have that for all values of k, p, and C,

$$\frac{\sqrt{k}}{\tau_S}\left(W_{\mathrm{FF2}_0}^* - W_{\mathrm{SF1}}^*\right) = \sqrt{e^{kp} - 1} - kp - 1 - C\ln\left(\frac{C}{1+C}\right) + \ln\left(\frac{2}{1+C}\right),$$

$$\frac{\sqrt{k}}{\tau_S}\frac{\partial\left(W_{\mathrm{FF2}_0}^* - W_{\mathrm{SF1}}^*\right)}{\partial p} = \frac{k(\sqrt{e^{kp} - 1} - 1)^2}{2\sqrt{e^{kp} - 1}} > 0.$$

Case of FF2 vs SF1. For all values of k, p, and C satisfying the FF2 condition in Eq. (A3), $C^2 > e^{kp} - 1$,

$$\frac{\sqrt{k}}{\tau_S}(W_{\mathrm{FF2}}^* - W_{\mathrm{SF1}}^*) = C[1 - \ln(2)] + C\ln[1 + (e^{kp} - 1)/C^2] - kp - 1 - C\ln\left(\frac{C}{1+C}\right) + \ln\left(\frac{2}{1+C}\right),$$

$$\frac{\sqrt{k}}{\tau_S}\frac{\partial(W_{\mathrm{FF2}}^* - W_{\mathrm{SF1}}^*)}{\partial p} = \frac{Cke^{kp}}{C^2 + e^{kp} - 1} - k = k(1 - C)\frac{C - (e^{kp} - 1)}{C^2 + (e^{kp} - 1)} \geqslant k(1 - C)\frac{C^2 - (e^{kp} - 1)}{C^2 + (e^{kp} - 1)} > 0.$$

In the last line $C \geqslant C^2$ and the FF2 condition is used in the first and the last inequality, respectively.

b. The SF2 subdomain

As for SF4 below, we show that the curvature of $W_{\mathrm{FF}}^* - W_{\mathrm{SF}}^*$ is negative in p for all values of k, p, and C on the SF2 domain.
Case of $\mathrm{FF2}_0$ vs SF2. For all values of k, p, and C satisfying the $\mathrm{FF2}_0$ condition $kp \leqslant \ln(2)$,

$$\frac{\sqrt{k}}{\tau_S}(W_{\mathrm{FF20}}^* - W_{\mathrm{SF2}}^*) = Ck(1 - p) + \sqrt{e^{kp} - 1} - kp - [1 - \ln(2)] - \ln[1 + C^2(e^{k(1-p)} - 1)],$$

$$\frac{\sqrt{k}}{\tau_S}\frac{\partial^2(W_{\mathrm{FF20}}^* - W_{\mathrm{SF2}}^*)}{\partial p^2} = -\frac{k^2 e^{kp}(2 - e^{kp})}{4(e^{kp} - 1)^{3/2}} - \frac{k^2 C^2(1 - C^2)e^{-k(1-p)}}{[C^2 + e^{-k(1-p)}(1 - C^2)]^2} < 0,$$

where the $kp \leqslant \ln(2)$ condition is used in the last inequality.
Case of FF2 vs SF2. For all values of k, p, and C,

$$\frac{\sqrt{k}}{\tau_S}(W_{\mathrm{FF2}}^* - W_{\mathrm{SF2}}^*) = Ck - (1 + C)kp - (1 - C)[1 - \ln(2)] + C\ln[(C^2 + e^{kp} - 1)/C^2] - \ln[1 + C^2(e^{k(1-p)} - 1)],$$

$$\frac{\sqrt{k}}{\tau_S}\frac{\partial^2(W_{\mathrm{FF2}}^* - W_{\mathrm{SF2}}^*)}{\partial p^2} = -k^2 C(1 - C^2)\left[\frac{e^{-kp}}{[1 - e^{-kp}(1 - C^2)]^2} + \frac{Ce^{-k(1-p)}}{[C^2 + e^{-k(1-p)}(1 - C^2)]^2}\right] < 0.$$

c. The SF4 subdomain

As for SF2 above, we show that the curvature of the difference is negative for increasing p on the SF4 domain.

Case of FF2$_0$ vs SF4. For all values of k, p, and C satisfying the FF2$_0$ condition $kp \leqslant \ln(2)$,

$$\frac{\sqrt{k}}{\tau_S}(W^*_{\mathrm{FF20}} - W^*_{\mathrm{SF4}}) = Ck(1-p) + \sqrt{e^{kp}-1} - \sqrt{(e^{kp}-1) + C^2(e^k - e^{kp})},$$

$$\frac{\sqrt{k}}{\tau_S}\frac{\partial^2(W^*_{\mathrm{FF20}} - W^*_{\mathrm{SF4}})}{\partial p^2} = -\frac{k^2 e^{kp}(2-e^{kp})}{4(e^{kp}-1)^{3/2}} + \frac{k^2 e^{kp}(1-C^2)}{4}\frac{(2-e^{kp}) - C^2(2e^k - e^{kp})}{[e^{kp}-1+C^2(e^k-e^{kp})]^{3/2}}$$

$$< -\frac{k^2 e^{kp}(1-C^2)}{4}\frac{C^2(2e^k - e^{kp})}{[e^{kp}-1+C^2(e^k-e^{kp})]^{3/2}} < 0,$$

where the $kp \leqslant \ln(2)$ condition is used in the first inequality in the last line.

Case of FF2 vs SF4. For all values of k, p, and C satisfying the FF2 condition $kp \leqslant \ln(2)$,

$$\frac{\sqrt{k}}{\tau_S}(W^*_{\mathrm{FF2}} - W^*_{\mathrm{SF4}}) = Ck(1-p) + C[1 - \ln(2)] + C\ln[1 + (e^{kp}-1)/C^2] - \sqrt{(e^{kp}-1) + C^2(e^k - e^{kp})},$$

$$\frac{\sqrt{k}}{\tau_S}\frac{\partial^2(W^*_{\mathrm{FF2}} - W^*_{\mathrm{SF4}})}{\partial p^2} = -\frac{k^2 e^{kp}(1-C^2)}{C^3}\left[\frac{1}{\left(\frac{e^{kp}-1}{C^2}+1\right)^2} - \frac{(1-C^2)(2-e^{kp}) - 2C^2(e^k-1)}{4[(1-C^2)\frac{e^{kp}-1}{C^2} + (e^k-1)]^{3/2}}\right]$$

$$\leqslant -\frac{k^2 e^{kp}(1-C^2)}{C^3}\left[\frac{1}{4} - \frac{(1-C^2)(2-e^{kp})}{4(e^k-1)}\right]$$

$$= -\frac{k^2 e^{kp}(1-C^2)}{4C^3}\left[\frac{(e^k-1) - (1-C^2)[1-(e^{kp}-1)]}{e^k-1}\right] < 0.$$

In the first inequality we use the FF2 condition $C^2 > e^{kp} - 1$ in the first fraction. In the second fraction, we first remove the negative term in the numerator, then the exponent in the denominator is set to 1 [the denominator is larger than 1 since $k > \ln(2)$], and finally the first positive term in the denominator is removed. In the last inequality, we use that $k > \ln(2)$ and $(1-C^2)[1 - (e^{kp}-1)] < 1$.

d. The SF3 subdomain

Set $y \equiv e^{kp} - 1 \in (0, 1)$, since $kp \leqslant \ln(2)$. Set $r \equiv C/\sqrt{(e^{kp}-1)(1-C^2)} = C/\sqrt{y(1-C^2)} > 0$. Set $d \equiv \sqrt{\delta^*_3}$ in Eq. (A2) such that

$$d = \frac{\sqrt{(e^{kp}-1)(1-C^2)} - C(e^{kp}-1)}{\sqrt{(e^{kp}-1)(1-C^2)} + C} = \frac{1 - ry}{1 + r} < \frac{1}{1+r}, \tag{B3}$$

where $d \in (0, 1)$ since $\delta^*_3 \in (0, 1)$. The inequality is due to $ry > 0$ and will be used later in the proof.

Case of FF2$_0$ vs SF3. The SF3 condition (vs SF1) $C < 2e^{-kp} - 1$ converts into $y < 1/(2r+1)$, while the FF2$_0$ condition $C \leqslant \sqrt{e^{kp}-1}$ in Eq. (A3) becomes $y \geqslant (r^2-1)/r^2$. This gives that $r \in (0, r^*)$, where $r^* \approx 1.19$ is the solution of $2r^3 - 2r - 1 = 0$. From Eqs. (A3) and (13),

$$\frac{\sqrt{k}}{\tau_S}(W^*_{\mathrm{FF20}} - W^*_{\mathrm{SF3}}) = \sqrt{e^{kp}-1} - \sqrt{\frac{(e^{kp}-1)(e^{kp}-1+d^2)}{e^{kp}-1+1}} - C[1 - \ln(2) + \ln(1-d) + d]$$

$$> \sqrt{y}\left(1 - \sqrt{\frac{y+d^2}{y+1}}\right) - C[1-\ln(2)] > \sqrt{y}\left(1 - \sqrt{\frac{y+\frac{1}{(1+r)^2}}{y+1}}\right) - \sqrt{y}\sqrt{1-C^2}\frac{C}{\sqrt{y(1-C^2)}}[1-\ln(2)]$$

$$> \sqrt{y}\left[\left(1 - \sqrt{\frac{y+\frac{1}{(1+r)^2}}{y+1}}\right) - r[1-\ln(2)]\right] \equiv \sqrt{y}H(r, y).$$

In the first inequality we use that $\ln(1-d) + d < 0$ for $d \in (0, 1)$ (easily shown by differentiation). In the second inequality, we use that $d < 1/(1+r)$ in Eq. (B3). In the third inequality, we use that $\sqrt{1-C^2} < 1$ in the last term.

We now show that for fixed r, $H(r, y)$ is decreasing when y increases from $y = 0$ towards the SF3-SF1 border where $y = 1/(2r+1)$. We then show that $H(r, y) > 0$ on the SF3-SF1 border, which means that $H(r, y) > 0$ on the whole SF3 $\cap$ FF2$_0$

domain

$$\frac{\partial H}{\partial y} = -\frac{1}{2}\sqrt{\frac{y+1}{y+a^2}}\frac{(1-a^2)}{(y+1)^2} < 0,$$

where $a \equiv 1/(1+r) < 1$. At the SF3-SF1 border,

$$H(r, y = 1/(2r+1)) = 1 - \sqrt{\frac{2+4r+r^2}{2(1+r)^3}} - r[1 - \ln(2)] \equiv h(r)$$

and

$$h'(r) = \frac{1}{2}\sqrt{\frac{1+r}{1+2r+0.5r^2}}\frac{1+3r+0.5r^2}{1+3r+3r^2+r^3} - [1 - \ln(2)].$$

Both denominators are increasing faster in r than their respective numerators, which gives that $h''(r) < 0$. Moreover, $h(0) = 0$ and $h(1.20) = 0.0097 > 0$. This means that $h(r) > 0$ on the whole range of $r \in (0, r^*)$.

Case of FF2 vs SF3. Set $R \equiv \sqrt{y}/C = \sqrt{e^{kp}-1}/C > 0$ such that $r = 1/\sqrt{R^2 - y}$. The FF2 condition $C > \sqrt{e^{kp}-1}$ in Eq. (A3) converts into $R < 1$, while the SF3 condition $C < 2e^{-kp} - 1$ becomes $R > \sqrt{y}(1+y)/(1-y)$. This gives that $y \in (0, y^*)$, where $y^* \approx 0.30$ is the solution of $y^3 + y^2 + 3y - 1 = 0$.

From Eqs. (A3) and (13),

$$\frac{\sqrt{k}}{\tau_S}(W^*_{\text{FF2}} - W^*_{\text{SF3}}) = C\ln\left(1 + \frac{e^{kp}-1}{C^2}\right) - \sqrt{\frac{(e^{kp}-1)(e^{kp}-1+d^2)}{e^{kp}-1+1}} - C[d + \ln(1-d)]$$

$$> C\ln\left(1 + \frac{y}{C^2}\right) - \sqrt{y}\sqrt{\frac{y+d^2}{y+1}} > C\ln(1+R^2) - C\frac{\sqrt{y}}{C}\sqrt{\frac{y+\frac{1}{(1+r)^2}}{y+1}}$$

$$= C\left(\ln(1+R^2) - R\sqrt{\frac{y+\frac{1}{[1+(R^2-y)^{-0.5}]^2}}{y+1}}\right) \equiv CG(R, y).$$

In the first inequality we use that $\ln(1-d) + d < 0$ for $d \in (0, 1)$. In the second inequality, we use that $d < 1/(1+r)$ in Eq. (B3).

We now show that for fixed R, $G(R, y)$ is decreasing when y increases from $y = 0$ towards the SF3-SF1 border where $R = \sqrt{y}(1+y)/(1-y)$. We then show that $G(R, y) > 0$ on the SF3-SF1 border, which means that $G(R, y) > 0$ on the whole SF3 $\cap$ FF2 domain.

Set $K(R, y) \equiv \sqrt{R^2 - y}$. Then the expression inside the square root in $G(R, y)$ can be written

$$F(y, K) \equiv \frac{y + (1 + \frac{1}{K})^{-2}}{y+1}, \quad \frac{dF}{dy} = \frac{\partial F}{\partial y} + \frac{\partial F}{\partial K}\frac{\partial K}{\partial y} = \frac{(K+1)(2K+1) - \sqrt{y}(y+1)}{(y+1)^2(K+1)^3} > 0.$$

The inequality is due to $K + 1 > 1 > \sqrt{y}$ and $2K + 1 > y + 1$ in the nominator. The latter stems from the SF3 condition:

$$K^2 = R^2 - y > \frac{y(1+y)^2}{(1-y)^2} - y > y(1+y) - y = y^2.$$

Since $\frac{dF}{dy} > 0$,

$$\frac{\partial G}{\partial y} = -R\frac{\frac{dF}{dy}}{2\sqrt{F(R, y)}} < 0.$$

It remains to show that $G(R, y) > 0$ on the SF3-SF1 border

$$G\left(R = \frac{\sqrt{y}(1+y)}{(1-y)}, y\right) = \ln(1 + R^2) - \frac{y}{1-y}\sqrt{1 + y + \frac{4y}{y+1}} \equiv g(y),$$

where $y \in [0, y^*]$ and $y^* \approx 0.3$. The condition $g(y) \geqslant 0$ is equivalent to $f(y) \equiv \exp[g(y)] - 1 \geqslant 0$. We use $e^{-x} \geqslant 1 - x + \frac{x^2}{2} - \frac{x^3}{6}$, which yields

$$f(y) \geqslant R^2 + (1+R^2)\left(-\frac{y\sqrt{Q(y)}}{1-y} + \frac{1}{2}\frac{y^2 Q(y)}{(1-y)^2} - \frac{1}{6}\frac{y^3 Q(y)^{3/2}}{(1-y)^3}\right) \geqslant R^2 + (1+R^2)\left(-\frac{y\sqrt{Q(y)}}{1-y} + \frac{1}{2}\frac{y^2 Q(y)}{(1-y)^2}(1 - \mathcal{A}y)\right),$$

(B4)

where $Q(y) \equiv 1 + y + \frac{4y}{1+y}$ and $\mathcal{A} \equiv \frac{\sqrt{Q(y^*)}}{3(1-y^*)} = \max(\frac{\sqrt{Q(y)}}{3(1-y)}) \approx 0.71$. Inserting $R^2 = y(\frac{1+y}{1-y})^2$ into Eq. (B4) and multiplying by $(1-y)^4(1+y)/y$, we obtain the condition

$$(1+y)^3(1-y)^2 + (1-y+3y^2+y^3)\left[(y-1)\sqrt{(1+y)^3 + 4y(1+y)} + \frac{y}{2}(1+6y+y^2)(1-\mathcal{A}y)\right] \geqslant 0$$

for $f(y) \geqslant 0$. We move the term with the square root to the right-hand side and then take the square of the expressions on both sides of the inequality. Since $1 - \mathcal{A}y > 0$ and $1 - y + 3y^2 + y^3 = (1+R^2)(1-y)^2 > 0$, the resulting inequality is the sufficient condition for $f(y) \geqslant 0$. It can be written as $\phi(y) \geqslant 0$, where

$$\phi(y) \equiv \left[(1+y)^3(1-y)^2 + \frac{y}{2}(1-y+3y^2+y^3)(1+6y+y^2)(1-\mathcal{A}y)\right]^2 - (1-y)^2(1-y+3y^2+y^3)^2(1+7y+7y^2+y^3)$$

$$= \left[1 + \frac{3y}{2} + \frac{1-\mathcal{A}}{2}y^2 - \left(3 + \frac{5\mathcal{A}}{2}\right)y^3 + (10+\mathcal{A})y^4 + \left(\frac{11}{2} - 9\mathcal{A}\right)y^5 + \frac{1-9\mathcal{A}}{2}y^6 - \frac{\mathcal{A}y^7}{2}\right]^2$$

$$- (1 - 4y + 12y^2 - 20y^3 + 22y^4 - 12y^5 - 4y^6 + 4y^7 + y^8)(1 + 7y + 7y^2 + y^3) \tag{B5}$$

$$\geqslant \left[1 + \frac{3y}{2} + ay^2 - by^3 + cy^4\right]^2 - (1 - 4y + 12y^2 - 20y^3 + 22y^4)(1 + 7y + 7y^2 + y^3), \tag{B6}$$

where $a \equiv \frac{1}{2}(1-\mathcal{A})$, $b \equiv 3 + \frac{5\mathcal{A}}{2}$, and $c \equiv 10 + \mathcal{A} + (\frac{11}{2} - 9\mathcal{A})y^* + \frac{1}{2}(1-9\mathcal{A})y^{*2} - \frac{1}{2}\mathcal{A}y^{*3}$. The inequality in Eq. (B6) is obtained replacing y^{4+n} by $y^{*n}y^4$ for $n = 1, 2, 3$ in square brackets of Eq. (B5) (the corresponding coefficients of these terms are negative and the resulting expression in the brackets becomes smaller but still positive since $a > 0$, $b > 0$, $c > 0$, and $by^3 < 1$), as well as omitting the terms $-12y^5 + 4y^7 = 4y^5(-3 + y^2) < 0$ and $-4y^6 + y^8 = y^6(-4 + y^2) < 0$ in the following brackets.

Further exact calculation of Eq. (B6) yields

$$\phi(y) \geqslant y^2[A_0 - A_1 z + A_2 z^2 + A_3 z^3 - A_4 z^4 - A_5 z^5 + c^2 y^6]$$
$$\geqslant y^2[A_0 - A_1 z + Bz^2] = y^2 \varphi(z), \tag{B7}$$

where $z = y/y^*$, $A_0 \equiv \frac{45}{4} + 2a$, $A_1 \equiv (37 + 2b - 3a)y^*$, $A_2 \equiv (a^2 - 3b + 2c + 38)y^{*2}$, $A_3 \equiv (3c - 2ab - 26)y^{*3}$, $A_4 \equiv (134 - b^2 - 2ac)y^{*4}$, $A_5 \equiv (22 + 2bc)y^{*5}$, and $B \equiv A_2 - A_4 - A_5$. The inequality (B7) holds for $y^* = 0.3$ because the coefficients A_n are positive in this case and $z \leqslant 1$. The values of coefficients at $y^* = 0.3$ are $a \approx 0.145$, $b \approx 4.775$, $c \approx 10.1909$, $A_0 \approx 11.54$, $A_1 \approx 13.8345$, $A_2 \approx 3.967$, $A_3 \approx 0.0861$, $A_4 \approx 0.8768$, $A_5 \approx 0.29$, and $B \approx 2.8003$.

The function $\varphi(z)$ has a single extremum (minimum) at $z = \frac{A_1}{2B} \approx 2.47 \notin [0, 1]$, so it changes monotonically from $\varphi(0) = A_0 \approx 11.54$ to $\varphi(1) = A_0 - A_1 + B \approx 0.5058$ within $z \in [0, 1]$, implying that $\varphi(z) > 0$ holds within this interval. Hence, $g(y) \geqslant 0$ holds for $y \in [0, y^*]$.

4. Proof of maximum relative distance between fast-first and slow-first policies

We propose that the maximum relative distance $(W_{FF}^* - W_{SF}^*)/W_{SF}^* = W_{FF}^*/W_{SF}^* - 1$ is in the intersection of the SF4 and the FF2$_0$ region. We therefore seek the maximum of

$$\frac{W_{FF2_0}^*}{W_{SF4}^*} = \frac{\sqrt{e^{kp} - 1} + Ck(1-p)}{\sqrt{(e^{kp} - 1) + C^2(e^k - e^{kp})}}$$

$$= \frac{\sqrt{e^{kp} - 1} + \sqrt{C^2(e^k - e^{kp})}\frac{k(1-p)}{\sqrt{(e^k - e^{kp})}}}{\sqrt{(e^{kp} - 1) + C^2(e^k - e^{kp})}}$$

$$= \frac{x + yd}{\sqrt{x^2 + y^2}} \equiv g(x, y, d),$$

where

$$x \equiv \sqrt{e^{kp} - 1}, \quad y \equiv \sqrt{C^2(e^k - e^{kp})}, \quad d \equiv \frac{k(1-p)}{\sqrt{e^k - e^{kp}}}. \tag{B8}$$

For fixed d, $g(x, y, d)$ is maximized when $y = dx$ (giving $\frac{\partial g}{\partial x} = 0 = \frac{\partial g}{\partial y}$). This gives $g(x, y = dx, d) = \sqrt{1 + d^2}$, which is maximized by maximizing d. For fixed k, $d = d(k, p)$ is decreasing in p since, when $kp < \ln(2)$,

$$\frac{\partial d}{\partial p} = -\frac{k}{2(e^k - e^{kp})^{3/2}}[2(e^k - e^{kp}) - e^{kp}k(1-p)]$$

$$< -\frac{k}{2(e^k - e^{kp})^{3/2}}[2(e^k - k) - 2(e^{kp} - kp)] < 0.$$

However, $p = 0$ implies that $x = 0 = y$, which is not in the domain of g. This means that p should be small, but positive. For fixed p, d is maximized by solving

$$\frac{\partial d}{\partial k} = \frac{(1-p)}{2(e^k - e^{kp})^{3/2}}[(2-k)e^k - (2 - kp)e^{kp}] = 0. \tag{B9}$$

This equation has only one solution $k = k(p)$, when we are restricted to $k > 0$. For $p = 0$, $k(0) = k^* \approx 1.5936$. Moreover, $k(p)$ is decreasing in p towards $k = 1$ when $p \to 1$. For small p, the solution of Eq. (B9) can be approximated by $k(p) \approx k^* - pk^*/(e^{k^*} - 2) \approx k^* - 0.55p$. In the limit when
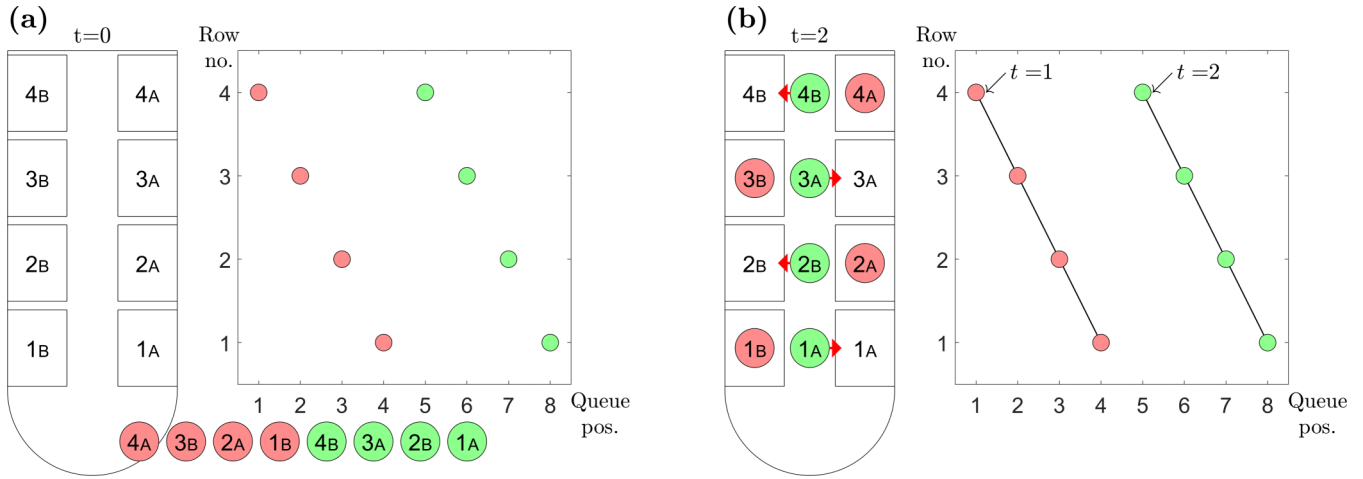
FIG. 13. (a) Initial position of the optimal queue ordering, with $N = 8$ passengers, 4 rows, and 2 seats per row. (b) The symmetry of the optimal queue ordering enables four passengers to sit down in each wave, which gives a minimal boarding time of $T = 2$. The time step $t = 1$ is not explicitly shown.

$p \to 0$, the maximum value of g is given by

$$\frac{W^*_{FF2_0}}{W^*_{SF4}} = \sqrt{1 + d^2(k^*, p)} = \sqrt{1 + \frac{k^{*2}(1-p)^2}{e^{k^*} - e^{k^* p}}}$$

$$\xrightarrow{p \to 0} \sqrt{1 + \frac{k^{*2}}{e^{k^*} - 1}} \approx 1.283\,59.$$

The value of C giving this maximum is found by setting $y = dx$ and using the expressions in Eq. (B8), which gives

$$C = \frac{k(1-p)}{e^k - e^{kp}} \sqrt{e^{kp} - 1} \approx \frac{k^{3/2}}{e^k - 1} \sqrt{p} \approx 0.513 \sqrt{p}.$$

The approximation holds for small p, and $k = k^*$ in the preceding expression. The solution $(k, p, C) = (k^*, p, 0.513\sqrt{p})$, where $p \approx 0$, is within the domains of $FF2_0$ and SF4. Theoretically, there could be other local maxima in other subdomains. However, numerical inspections indicate that the given solution is the global maximum.

APPENDIX C: REORGANIZED QUEUE WITH MINIMAL BOARDING TIME

In Fig. 13 the queue in Fig. 1 is reorganized to obtain a minimal boarding time.

[1] E. Bachmat, D. Berend, L. Sapir, S. Skiena, and N. Stolyarov, Analysis of aeroplane boarding via spacetime geometry and random matrix theory, J. Phys. A: Math. Gen. **39**, L453 (2006).

[2] V. Frette and P. C. Hemmer, Time needed to board an airplane: A power law and the structure behind it, Phys. Rev. E **85**, 011130 (2012).

[3] N. Bernstein, Comment on "Time needed to board an airplane: A power law and the structure behind it", Phys. Rev. E **86**, 023101 (2012).

[4] E. Bachmat, V. Khachaturov, and R. Kuperman, Optimal back-to-front airplane boarding, Phys. Rev. E **87**, 062805 (2013).

[5] Y. Baek, M. Ha, and H. Jeong, Impact of sequential disorder on the scaling behavior of airplane boarding time, Phys. Rev. E **87**, 052803 (2013).

[6] M. Brics, J. Kaupužs, and R. Mahnke, Scaling behavior of an airplane-boarding model, Phys. Rev. E **87**, 042117 (2013).

[7] E. Bachmat, Airplane boarding, disk scheduling, and Lorentzian geometry, *Mathematical Adventures in Performance Analysis: From Storage Systems, Through Airplane Boarding, to Express Line Queues* (Birkhäuser, Cham, 2014), Chap. 3, pp. 51–129.

[8] R. Mahnke, J. Kaupužs, and M. Brics, Air traffic, boarding and scaling exponents, in *Proceedings of Traffic and Granular Flow '13*, edited by M. Chraibi, M. Boltes, A. Schadschneider, and A. Seyfried (Springer International Publishing, Cham, 2015), pp. 305–314.

[9] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Space-Time as a Causal Set, Phys. Rev. Lett. **59**, 521 (1987).

[10] J. Myrheim, Statistical geometry, CERN Report No. CERN-TH-2538, 1978 (unpublished).

[11] G. t'Hooft, Quantum gravity: A fundamental problem and some radical ideas, in *Recent Developments in Gravitation*, edited by M. Lévy and S. Deser, NATO Advanced Studies Institute, Series B: Physics Vol. 44 (Springer, Boston, MA, 1979), pp. 323–345.

[12] G. Brightwell and R. Gregory, Structure of Random Discrete Spacetime, Phys. Rev. Lett. **66**, 260 (1991).

[13] F. Jaehn and S. Neumann, Airplane boarding, Eur. J. Oper. Res. **244**, 339 (2015).

[14] S. Neumann, Is the boarding process on the critical path of the airplane turn-around? Eur. J. Oper. Res. **277**, 128 (2019).

[15] T. Reed, Will American's new boarding process work? It failed at Virgin America, available at https://www.forbes.com/sites/tedreed/2013/05/18/will-americans-new-boarding-process-work-it-failed-at-virgin-america/#3dc5b919480c (2013).

[16] J. Audenaert, K. Verbeeck, and G. V. Berghe, Multi-agent based simulation for boarding, in *Proceedings of The 21st Benelux Conference on Artificial Intelligence, BNAIC 2009, Eindhoven, the Netherlands, 2009* (Eindhoven University of Technology, Eindhoven, 2009), pp. 3–10.

[17] R. J. Milne and A. R. Kelly, A new method for boarding passengers onto an airplane, J. Air Transp. Manag. **34**, 93 (2014).

[18] S.-J. Qiang, B. Jia, D.-F. Xie, and Z.-Y. Gao, Reducing airplane boarding time by accounting for passengers' individual properties: A simulation based on cellular automaton, J. Air Transp. Manag. **40**, 42 (2014).

[19] G. Notomista, M. Selvaggio, F. Sbrizzi, G. D. Maio, S. Grazioso, and M. Botsch, A fast airplane boarding strategy using online seat assignment based on passenger classification, J. Air Transp. Manag. **53**, 140 (2016).

[20] E. Bachmat, Airplane boarding meets express line queues, Eur. J. Oper. Res. **275**, 1165 (2019).

[21] E. Bachmat, D. Berend, L. Sapir, S. Skiena, and N. Stolyarov, Analysis of airplane boarding times, Oper. Res. **57**, 499 (2009).

[22] A. M. Vershik and S. V. Kerov, Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables, Dokl. Akad. Nauk SSSR **233**, 1024 (1977) [Sov. Math. Dokl. **18**, 527 (1977)].

[23] B. Logan and L. Shepp, A variational problem for random Young tableaux, Adv. Math. **26**, 206 (1977).

[24] J. Deuschel and O. Zeitouni, Limiting curves for i.i.d. records, Ann. Probab. **23**, 852 (1995).

[25] L. Hutter, F. Jaehn, and S. Neumann, Influencing factors on airplane boarding times, Omega **87**, 177 (2019).