Discrete-time TASEP with holdback

Seva Shneer[a], Alexander Stolyar[b],*

[a] *Heriot-Watt University, United Kingdom*
[b] *University of Illinois at Urbana-Champaign, United States of America*

Abstract

We study the following interacting particle system. There are ρn particles, $\rho < 1$, moving clockwise ("right"), in discrete time, on n sites arranged in a circle. Each site may contain at most one particle. At each time, a particle may move to the right-neighbor site according to the following rules. If its right-neighbor site is occupied by another particle, the particle does not move. If the particle has unoccupied sites ("holes") as neighbors on both sides, it moves right with probability 1. If the particle has a hole as the right-neighbor and an occupied site as the left-neighbor, it moves right with probability $0 < p < 1$. (We refer to the latter rule as a "holdback" property.) From the point of view of holes moving counter-clockwise, this is a zero-range process.

The main question we address is: what is the system steady-state flux (or throughput) when n is large, as a function of density ρ? The most interesting range of densities is $0 \le \rho < 1/2$. We define the system *typical flux* as the limit in $n \to \infty$ of the steady-state flux in a system subject to additional random perturbations, when the perturbation rate vanishes. Our main results show that: (a) the typical flux is different from the formal flux, defined as the limit in $n \to \infty$ of the steady-state flux in the system without perturbations, and (b) there is a phase transition at density $h = p/(1 + p)$. If $\rho < h$, the typical flux is equal to ρ, which coincides with the formal flux. If $\rho > h$, a *condensation* phenomenon occurs, namely the formation and persistence of large particle clusters; in particular, the typical flux in this case is $p(1 - \rho) < h < \rho$, which differs from the formal flux when $h < \rho < 1/2$.

Our results include both the steady-state analysis (which determines the typical flux) and the transient analysis. In particular, we derive a version of the Ballot Theorem, and show that the key "reason" for large cluster formation for densities $\rho > h$ is described by this theorem.

---

* Corresponding author.
  E-mail addresses: V.Shneer@hw.ac.uk (S. Shneer), stolyar@illinois.edu (A. Stolyar).

1. Introduction

1.1. The model and motivation

The basic model, which is the focus of this paper, is the following interacting particle system. There are n sites (or nodes), arranged in a circle and numbered from 0 to $n-1$ in the clockwise order. There is a constant number ρn of particles in the system, where $\rho \in (0, 1]$ is the particle density. There is at most one particle at each site at any time. The system evolves in discrete time $t = 0, 1, 2, \ldots$. The state at a given time is described as a sequence of particles (occupied sites) and holes (empty sites). We often refer to the clockwise and counter-clockwise directions as "right" and "left", respectively. The particles never move counter-clockwise. Their clockwise movement, at each time, is governed by the following rules:

(a) if a particle has another particle as a right-neighbor, it does not move;
(b) if a particle has holes as neighbors on both sides, it moves to the right-neighbor site with probability 1;
(c) if a particle has a hole as right-neighbor and a particle as left-neighbor, it moves to the right-neighbor site with fixed probability $p \in (0, 1]$.

This model may be considered as a version of the discrete-time Totally Asymmetric Simple Exclusion Process (TASEP), with parallel updates, cf. [7]. We refer to rule (c) as a "holdback" property and thus call the model discrete-time TASEP-H, where H stands for holdback. In the classical version of discrete-time TASEP a particle cannot move if its right-neighbor is another particle and, otherwise, moves right with a certain fixed probability. (For a general introduction into interacting particle systems see [9].)

This interacting particle system is motivated, in particular, by a simple model of packet movement in a wireless communication system under a CSMA (Carrier-Sense Multiple Access) protocol. Sites correspond to network nodes and particles correspond to data packets. Discrete time corresponds to the sequence of fixed-length time slots, in which packet transmissions occur. Each node can hold at most one packet. Packets "move" along a sequence of nodes, in the "right" direction, by being transmitted from nodes to their "right" neighbors. Consider a node that in a given time slot has a packet. If both neighbors of the node do not have packets, the packet will be transmitted successfully, because the transmission does not experience an interference (from neighbors). If the right-neighbor node is occupied by another packet, the packet will be certainly blocked from moving right. If the right-neighbor node is empty, but the left-neighbor node is occupied, the transmission protocol is such that the "competition" from the left-neighbor packet may prevent the packet from moving right, with some probability. (This motivating model is very basic. A more complicated – and more realistic – CSMA model, which exhibits a qualitatively similar behavior, is discussed in Appendix B.) Our TASEP-H system is also motivated by and related to a slow-to-start model [4] of car traffic, where particles represent cars and the holdback property corresponds to the fact that cars need some time to accelerate after being stopped. The relation of TASEP-H to the slow-to-start model will be discussed in more detail in Section 6.2.

Going back to the TASEP-H system definition, note that from the point of view of holes moving counter-clockwise (left), the rules (a)–(c) are equivalent to the following ones:

(a′) if a hole has another hole immediately to the left, it does not move;
(b′) if a hole has a particle immediately to the left, followed by another hole, it moves to the left-neighbor site;

(c′) if a hole has two or more particles immediately to the left, it moves to the left-neighbor site with probability $p \in (0, 1]$.

In the latter interpretation (as holes' movement counter-clockwise), TASEP-H system is a special case of the model considered in [7] where the probability of a hole movement may be an arbitrary function of the number of holes immediately to the left of it. It is shown (see [7, Section 7]) that the stationary distribution of the system may be presented in a product form.

In general, the class of TASEP models is very rich and has received a lot of attention in the literature, because it has many applications, including statistical physics and transportation; see, e.g.[1,5,8,9]. In addition to [7], our model is related to some other TASEP models considered in the literature. We mention q-TASEP [3] for the continuous-time and [2] for the discrete-time version; both papers however consider movement of a finite number of particles on an infinite line.

1.2. Particle flux: "formal" vs. "typical"

The particle *flux* ϕ is defined as a steady-state average number of particle movements, per site per time unit. (This corresponds to the throughput in the context of wireless systems.) In this paper we are primarily interested in the dependence of flux ϕ on the density ρ, when n is large.

Trivially, for any n, when $\rho < 1/2$, the flux is simply equal to the density, $\phi = \rho$, because the system eventually enters a *completely sparse* state, where all particles are *free* (have holes as both neighbors), and all particles will move at speed 1 thereafter. Following the discussion which will be given shortly, it is also not hard to guess (and observe in simulations, and prove) that when $\rho > 1/2$, the steady-state of the system is very different — the system spends most of the time in a *condensed* state, where a large particle cluster (a sequence of contiguously occupied sites) exists. Correspondingly, the "formal" flux of TASEP-H is as follows: $\phi = \rho$ for $\rho < 1/2$, but it has a discontinuity ("negative jump") at $1/2$ and is continuous decreasing for $\rho > 1/2$. Fig. 1 shows this "formal" flux, when n is large.

However, an interesting phenomenon can be observed in simulations. If $h = p/(1 + p) < \rho < 1/2$ and n is large, and we start the process in a state chosen uniformly at random, then the process enters a quasi-stationary regime, in which it stays in condensed states (with one or more clusters). This quasi-stationary regime persists for a very long time. The flux corresponding to this quasi-steady-state is strictly less than ρ. If we consider as "typical" the system flux that is observed with very high probability for a very long time, then this "typical" flux of TASEP-H is shown in Fig. 1 (for large n): it depends on ρ continuously, coincides with "formal" flux for $\rho < h = p/(1 + p)$ and $\rho > 1/2$, but is different for $h < \rho < 1/2$.

In this paper we will formally define the "typical" flux as the limit in $n \to \infty$ of the flux of the TASEP-H system subject to additional state perturbations, as the perturbation rate (per particle per time unit) vanishes as $n \to \infty$. Our results will show, in particular, that the typical flux of TASEP-H is indeed as shown in Fig. 1. Most importantly, for $h < \rho < 1/2$, the typical flux of TASEP-H is equal to that in quasi-stationary regime (as opposed to the formal flux, which is ρ).

1.3. Condensation. Quasi-stationary regime

The term *condensation* in interacting particle systems is inspired by the famous Bose condensation [6,7]. (Regarding the terminology, in most of the literature, the condensation occurs
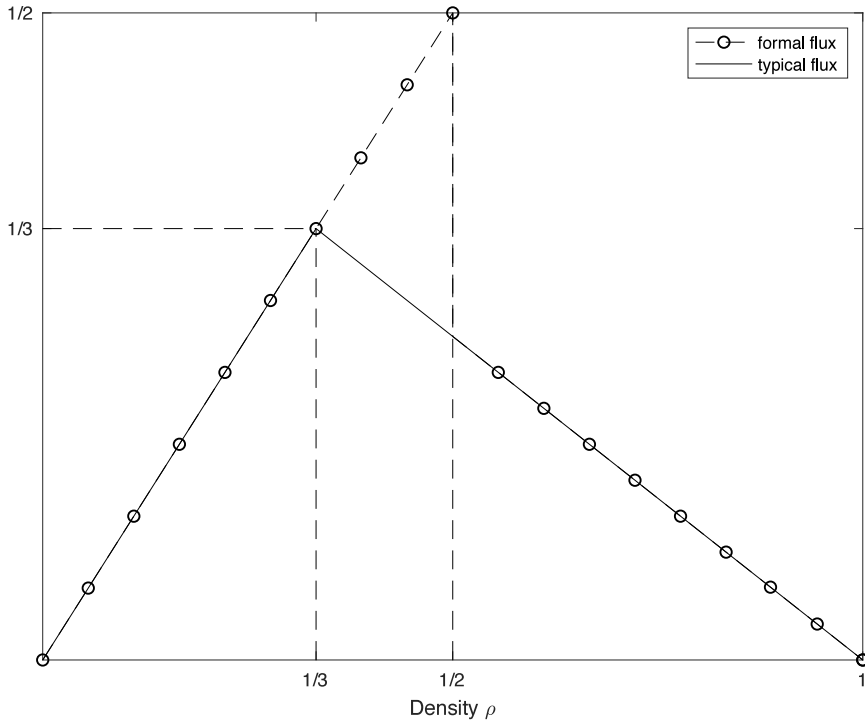
Fig. 1. Large n; $p = 1/2$. Formal and typical flux versus ρ. For $\rho < 1/3$ and $\rho > 1/2$, the formal and typical fluxes coincide.

at *low* particle density, as opposed to high. This is consistent with TASEP-H condensation, if we view the process from the point of view of holes' movement.)

We now discuss the condensation effect in more detail. Recall that the state of the system at any time is described as a sequence of particles and holes. We will refer to a contiguous segment of (at least two) particles as a cluster. Only the particle at the right end of a cluster may move at a given time. We will also refer to a contiguous segment of sites (which may include all sites) as a sparse interval, if all particles in it are *free* (have holes as both neighbors). Obviously, each free particle at a given time moves right with probability 1.

A system state such that all particles are free we will call (interchangeably) *completely sparse*, or *ideal*, or *absorbing*. As discussed earlier, if $\rho < 1/2$, then for any n and any p, the system eventually reaches an absorbing state and therefore its flux $\phi = \rho$.

However, for $h < \rho < 1/2$, it may take the system a very long time to reach such an absorbing state. Instead, the system may enter a *quasi-stationary* regime, which will persist for a long time. (It is reasonable to guess that this time is exponential in n, but we will not need this estimate in this paper.) We explain this – heuristically at this point – as follows. (Our main results will formally substantiate this heuristics.) Suppose, the system starts in a state such that $\tau^* n$ particles, $0 < \tau^* \leq 1$, form a single cluster, while the remaining sparse interval, consisting of $(1 - \tau^*)n$ sites, has particles spread out with the density

$$h = p/(1 + p).$$

Note that in this case the overall density of particles is $\rho = \tau^* + (1 - \tau^*)h > h$, and then

$$\tau^* = \rho(1 + p) - p.$$

Then, the system dynamics is as follows. The particles will leave the cluster from its "right end" at the rate p, with the average distance between the released particles being $1 + 1/p = 1/h$. (Indeed, when a particle becomes the right-end particle of a cluster, the random number X of time slots it takes for this particle to leave the cluster has geometric distribution, $\mathbb{P}\{X = i\} = p(1 - p)^{i-1}$, $i = 1, 2, \ldots$, with $\mathbb{E}X = 1/p$. Then X is also the number of holes between two consecutive released particles. Then, the average distance between two consecutive released particles is $1 + \mathbb{E}X = 1 + 1/p$.) Therefore, the cluster right end, as it moves left at the rate p, leaves particle density exactly h in "its wake" on the right. Of course, free particles (in the sparse interval) move right at speed 1. It is not hard to see that, as free particles "hit" (and thus join) the cluster on its left, the left end moves left at the rate p. (This follows because the average time between two consecutive free particles hitting the cluster on the left is equal to the average number of holes separating them, which is $1/p$.) To summarize, the cluster moves left at speed p, with its length $\tau^* n$ staying (approximately) constant, and with the density "everywhere" in the sparse interval staying equal to h. Therefore, the system state remains (almost) invariant, up to the shift of the cluster. We will call such a state a *main equilibrium state* (MES). (This definition will be made formal later in the paper.) Note that a MES exists and is unique (up to space shift) if and only if $\rho > h$. As long as a MES remains the system state, the flux is $\phi^* \doteq (1 - \tau^*)h = p(1 - \rho) < h < 1/2$. The existence of a quasi-stationary regime, represented in our model by MES, is often referred to as *metastability*.

Further notice that, if the system starts in a MES, it will take a long time (which, again, can be guessed to be exponential in n) for the cluster to "dissolve". Therefore, if the density $\rho \in (h, 1/2)$, the system may enter a MES and stay in it – as opposed to an absorbing state – for a very long time, during which it will have the flux $\phi^* < h < \rho$, as opposed to the "formal" flux ρ.

Note that MES exists for any $\rho > h$, not just $h < \rho < 1/2$. The graph of ϕ^* (when the system stays in a MES) versus ρ, for $\rho > h$, can be seen in Fig. 2, for the cases $p = 0.1$, $p = 0.5$ and $p = 0.9$. Note that ϕ^* is linearly decreasing from h to 0 in the interval $[h, 1]$. The case when $\rho \geq 1/2$ is a simplified version of the case $h < \rho < 1/2$. For this reason, in this paper we restrict our attention to the most interesting case $\rho < 1/2$.

1.4. Main results and contributions

To summarize the previous subsection, even when $\rho < 1/2$, in addition to a formal stationary regime, when the system stays in an absorbing state and the flux is ρ, there may exist a quasi-stationary regime (represented by a MES), in which the system flux is ϕ^*. This raises the following question. What should be considered a "typical" flux of TASEP-H? Is it ρ, or ϕ^*, or maybe something else? In this paper we introduce the notion of typical flux, which is, informally speaking, the value of the flux in the system subject to infrequent random perturbations of the system state, in addition to the system "normal" random evolution. This is motivated by the fact that, even if the basic system behavior is described by TASEP-H – whether it is movement of packets in a wireless system or movement of cars on a road – practical systems are subject to occasional "imperfections", or perturbations. For example, new packets [resp., cars] may occasionally enter or leave the system [resp., road]; packets may occasionally fail to transmit successfully even in the absence of interference; cars may break
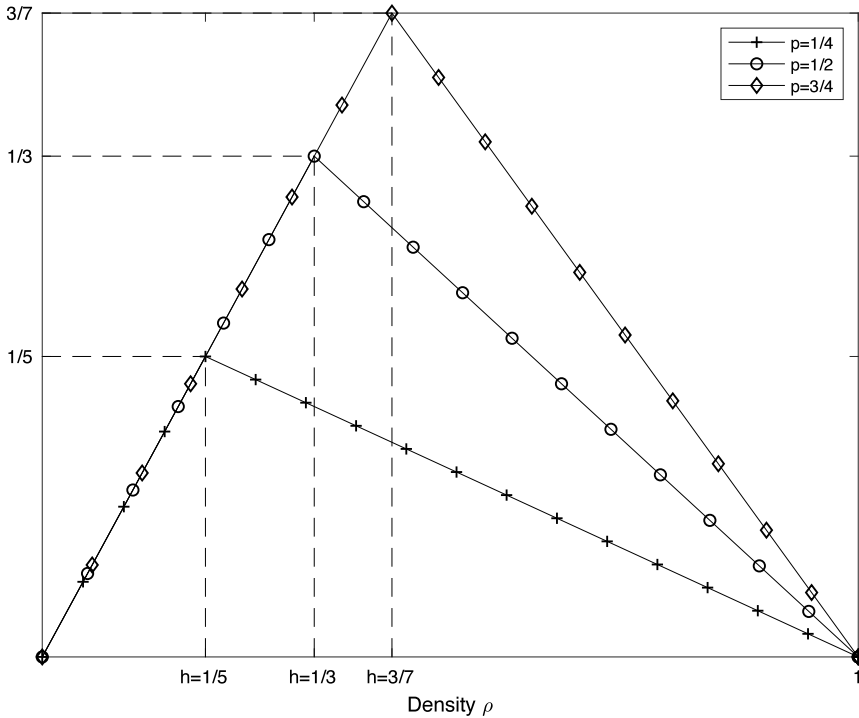
Fig. 2. Large n; $p = 1/4, 1/2, 3/4$. Typical flux versus density ρ.

down; etc. If occasional perturbations result in the flux that is different from the formal flux, then the former rather than latter can be considered typical. More specifically, in this paper, for a given density ρ, we will define the typical flux $\phi(\rho)$ as the limit in $n \to \infty$ of the flux of TASEP-H *with perturbations*, as the rate of perturbations (per particle per time unit) vanishes as $n \to \infty$. The main goal of this paper is to demonstrate that even very low perturbation rate, $O(1/n^2)$, leads to the typical flux given by Fig. 1.

Perturbations in the basic TASEP-H may be defined in different ways. For example, we may replace the probability 1 with which a free particle moves right, with a close to one probability $\pi < 1$. Such a system, which we will refer to as *zero-range* model, is still within the framework of [7], its stationary distribution has a product form, and the limiting flux ϕ_π, as $n \to \infty$, can be found. We in fact carry out this analysis in Section 5, and demonstrate that $\lim_{\pi \uparrow 1} \phi_\pi$ is exactly as depicted in Fig. 2. Note that here we take the limit $n \to \infty$ first, and limit $\pi \uparrow 1$ second. This means, in particular, that as $n \to \infty$, the perturbation rate remains "high", namely $O(1)$ per particle per time unit.

However, the above analysis, based on the steady-state product-form for the zero-range model has a number of limitations, including the following.

- It is interesting to look at how low a perturbation rate can be so that the flux remains as in Fig. 2. This question could be answered by considering a zero-range model with $\pi = \pi(n) \uparrow 1$. Unfortunately, analysis of the flux in this latter case (i.e., taking limit $n \to \infty$ with $\pi = \pi(n) \uparrow 1$), appears to be difficult.

- Perturbations induced by $\pi < 1$ are not the only type of perturbations of interest. And for other types of perturbations, as those we introduce shortly, the product-form for the steady-state may no longer be available.
- An analysis of steady-state alone does not shed light into how condensation occurs, i.e. into the questions of the following type. What is the mechanism and dynamics of large cluster formation? How fast do the clusters form? And so on.

To address these issues, most of the analysis in this paper concerns a different type of perturbations (introduced shortly) and "low" perturbation rates (per particle per time unit), vanishing as $n \to \infty$. For these perturbations the product form of the steady-state is *not* available. In particular, we do a transient analysis which sheds light on why, for the densities $\rho > h$, even a small cluster formed by a perturbation has a positive, uniformly bounded away from 0, probability to grow large, bring the system to a MES, which will then persist for a long time. The key "reason" for large cluster formation for densities $\rho > h$ is described by a version of the Ballot Theorem, which we derive in Proposition 8.

The TASEP-H with perturbations, which is the main focus of this paper, is as follows. We define two versions of the perturbation mechanism, both preventing the system from being absorbed in an ideal (absorbing) state. For both of them, a perturbation is defined as follows: a particle, chosen uniformly at random, is relocated to one of the holes, also chosen uniformly at random. The two versions of TASEP-H with perturbations are:

- *A-perturbations (perturbations of absorbing states).* At every time, in which the system enters an ideal (absorbing) state, a perturbation happens with probability λ/n with some $\lambda > 0$, independent of the process history up to that time.
- *I-perturbations (independent perturbations).* At every time a perturbation happens with probability λ/n with some $\lambda > 0$, independently of the process history up to that time.

We analyze the behavior of the system as $n \to \infty$, with ρ staying constant. In particular, we study the process under hydrodynamic space/time scaling. In addition to the steady-state results – in fact, as a tool for obtaining them – we derive results on the process transient behavior under this scaling. These transient results may be of independent interest.

Our **main results** are as follows.

- For **TASEP-H with A-perturbations** we prove that the limit (in $n \to \infty$) of stationary distributions is such that:
    - If $\rho < h$, a cluster (if any) contains zero fraction of sites; consequently, the typical flux $\phi(\rho) = \rho$; see Theorem 9.
    - If $\rho > h$, the system state is the MES; consequently, the typical flux $\phi(\rho) = \phi^* = p(1 - \rho)$; see Theorem 12.

- For **TASEP-H with I-perturbations** we prove that the limit (in $n \to \infty$) of stationary distributions is such that:
    - If $\rho < h$, all clusters (if any) contain zero fraction of sites; consequently, the typical flux $\phi(\rho) = \rho$; see Theorem 17.
    - If $\rho > h$, the typical flux $\phi(\rho) < \rho$; see Theorem 19.

We conjecture that, if $\rho > h$, the typical flux $\phi(\rho) = \phi^*$; Conjecture 18.

- For the **zero-range model** we prove that if we first take the limit of stationary distributions as $n \to \infty$, and then the limit in $\pi \uparrow 1$, the limit is exactly equal to the typical flux $\phi(\rho)$ of TASEP-H with A-perturbations; see Theorem 20 in Section 5.
- We derive a version of the **Ballot Theorem** (Proposition 8) which serves as a key tool for establishing condensation at high densities $\rho > h$.

1.5. Paper organization

The rest of the paper is organized as follows. Some basic notation, used throughout the paper, is given in Section 1.6. Section 2 defines the basic TASEP-H model (without perturbations), the model with perturbations, and a system flux; it also defines the asymptotic regime, the corresponding limiting flux (typical flux), the process under hydrodynamic scaling, and the probability space construction. The analysis of TASEP-H with A-perturbations is given in Section 3: we define the process hydrodynamic limits and establish their properties (Section 3.1); we state a version of the Ballot Theorem (Proposition 8 in Section 3.2); the main results for TASEP-H with A-perturbations (Theorems 9 and 12) are stated and proved in Section 3.3. Section 4 is devoted to the analysis of TASEP-H with I-perturbations: some preliminary facts are given in Section 4.1; hydrodynamic limits are defined and studied in Section 4.2; Sections 4.3 and 4.4 contain our main results and the conjecture (Theorems 17, 19, Conjecture 18) for TASEP-H with I-perturbations. In Section 5 we formally define the zero-range model and derive (Theorem 20) the limit of its flux (as $n \to \infty$ and then $\pi \to 1$). In Section 6 we present and discuss further conjectures for both the TASEP-H with perturbations and zero-range models. Appendix A contains the proof of Proposition 8; Appendix B discusses a CSMA model further (indirectly) motivating the holdback property of TASEP-H.

1.6. Basic notation

We denote by $\mathbb{R}$ and $\mathbb{Z}$ the sets of real numbers and integers, respectively, and by $\mathcal{L}$ the Lebesgue measure on $\mathbb{R}$. Abbreviation a.e. means *almost everywhere w.r.t. Lebesgue measure*. The minimum and maximum of two numbers are denoted $a \wedge b \doteq \min(a, b)$ and $a \vee b \doteq \max(a, b)$, respectively; $\lfloor a \rfloor$ is the largest integer which is $\leq a$; $\lceil a \rceil$ is the smallest integer which is $\geq a$. $\mathbb{I}(A)$ is the indicator of event or condition A. RHS and LHS mean *right-hand side* and *left-hand side*, respectively. WLOG means *without loss of generality*.

A function is called c-Lipschitz if it is Lipschitz continuous with constant $c \geq 0$. We say that a function is RCLL if it is *right-continuous with left limits*. Notation $(\partial^-/\partial x)f(x, t)$ means left partial derivative in x. U.o.c. convergence means *uniform on compact sets* convergence. Notation $O(1/n)$ signifies a function $g(n)$ such that $\limsup_{n \to \infty} |ng(n)| < \infty$.

Abbreviation *r.v.* means *random variable*; *w.p.1* means *with probability* 1; *i.i.d.* means *independent identically distributed*. Probability distributions are defined on the spaces and corresponding σ-algebras that are clear from the context. We denote by $\Rightarrow$ the convergence of random elements in distribution. For a random process $Y(t), t \geq 0$, we denote by $Y(\infty)$ its (random) state in a stationary regime. (In other words, the distribution of $Y(\infty)$ is a stationary distribution of $Y(t)$.) We will say that a random variable X has distribution $GEOM(\ell, p)$ [or simply write $X \sim GEOM(\ell, p)$], for integer ℓ and real $0 \leq p \leq 1$, if $\mathbb{P}\{X = i\} = p(1-p)^{i-\ell}$, $i \geq \ell$; $\mathbb{E}X = \ell + (1 - p)/p$.

2. Model

2.1. Basic model

Consider n sites arranged in a circle. The sites are numbered from 0 to $n - 1$ in clockwise order. There is a constant number ρn, $0 < \rho \leq 1$, of particles in the system, with either 0 or 1 particles located at each site at any time; correspondingly, a site may be empty (0 particles) or occupied (1 particle). We often refer to empty sites as holes, and to occupied sites as particles. The clockwise [resp., counterclockwise] direction we often refer to as "right" [resp. "left"] direction.

The system evolves in discrete time $t = 0, 1, 2, \ldots$. A time index t is sometimes referred to as time slot t. The system state at each time is its sites' occupancy configuration, i.e. the sequence of particles and holes. Given the state at time t, the (random) state at time $t + 1$ is determined by applying the following rules to each particle independently:

(a) if the right-neighbor of the particle is another particle, it does not move;

(b) if both the right-neighbor and left-neighbor of the particle are holes, the particle moves to the right-neighbor site with probability 1;

(c) if the right-neighbor of the particle is a hole and the left-neighbor is another particle, the particle moves to the right-neighbor site with probability $p \in (0, 1]$.

We refer to the rule (c) as a "holdback" property.

To make some definitions that follow later unambiguous, we adopt the following convention about the exact timing of the particles' movement. Namely, the movement of particles that changes the system state from that at $t - 1$ to that at t is attributed to time t.

We will use the following terminology throughout. We call a particle *free* if it has holes on both sides. A contiguous set of particles is a *cluster* if it has holes as neighbors on both sides. A contiguous set of sites is a *sparse interval*, if it contains only free particles and holes (i.e., does not overlap with any cluster). A system state is called *completely sparse*, or *ideal*, if it contains no clusters.

Note that any trajectory of the basic model is such that the number of clusters cannot increase. Indeed, free particles cannot form new clusters – they can only join existing ones. If we consider any cluster initially present, then as particles "leave" this cluster from the right and "join" it on the left, the cluster may "move" to the left, may grow or decrease in length, or may eventually disappear; but it cannot split into two or more clusters and, therefore, no new cluster can ever be created.

2.2. TASEP-H with perturbations

We will now introduce the TASEP-H model with perturbations, in fact two different versions of it. For a given system state its (random) *perturbation* is defined as follows: we pick a particle uniformly at random, remove it, and then place it into one of the holes picked uniformly at random. We adopt a convention that the site from which the particle was removed is immediately considered a hole, so that the particle may go back to it. (This convention is not essential.)

We also adopt the following convention about the exact timing of a perturbation with respect to time slots. Applying a perturbation at time t means that it is applied to the state the system enters after the "normal" particles' movement at time t (i.e. the movement "between" $t - 1$ and t), and it is done before the next time $t + 1$. In other words, the perturbation does not take

an extra time slot, and the final system state at time t is the state after the normal particles' movement at t and after a perturbation (if any) is applied.

Model with I-perturbations. In this version, at each time t, one perturbation is applied with probability λ/n, $\lambda > 0$, independently of the process history up to t. (Term I-perturbations is because the perturbations are independent.)

Model with A-perturbations. In this version, one perturbation is applied with probability λ/n, $\lambda > 0$, at each time t in which the system enters to – or stays at – an ideal state, *after the normal particles' movement.* (Term A-perturbations is because the perturbations explicitly prevent absorption.)

Recall that in the basic model (without perturbations) no new cluster can ever be created. Consequently, in a model with perturbations, *in the time intervals between perturbations* the number of clusters cannot increase. However, unlike in the basic model, a model with perturbations is such that new clusters may be created by perturbations. If a perturbation is applied to an ideal state, then at most one cluster (of length two or three) may be created. Note that for the A-perturbation model this necessarily means that, *after the system "hits" an ideal state,* it can have at most one cluster from that time on. In the I-perturbation model a perturbation may: (a) split a cluster into two clusters and/or (b) merge two clusters by filling a hole separating them and/or (c) create a new cluster (of length two or three).

For future reference, we summarize these simple observations as the following

Lemma 1. *Any process trajectory has the following properties, depending on the model type.*

(i) The basic model (without perturbations). The number of clusters cannot increase. Each initial cluster may move to the left, may grow or decrease in length, may eventually disappear. But no new cluster can ever be created.

(ii) A-perturbations model. The number of clusters cannot increase until a trajectory "hits" an ideal state. There can be at most one cluster from that time on, which may be created by a perturbation.

(iii) I-perturbations model. The number of clusters cannot increase between perturbations. Upon a perturbation, a new cluster may form and/or a cluster may split into two and/or two clusters may merge.

The process without perturbations behaves as follows. If $\rho \leq 1/2$, the behavior is rather trivial: an ideal state is eventually reached, and from that time on all particles remain free and move at speed 1. If $\rho > 1/2$: at least one cluster always exists; the states with more than one cluster are transient, so that a state with exactly one cluster is eventually reached; the process on the states with exactly one cluster is a discrete-time finite (aperiodic) Markov chain; therefore, the process has a unique stationary distribution.

For both the A-perturbation and I-perturbation models, the process is a discrete-time finite (aperiodic) Markov chain. For I-perturbations, the Markov chain is irreducible. For A-perturbations, the Markov chain on the set of states with no more than one cluster is irreducible; states with more than one cluster are transient. (So, as far as a stationary distribution of the process with A-perturbations is concerned, the states with more than one cluster can be ignored.) Therefore, for both perturbation models, the Markov chain has a unique stationary distribution.

2.3. Flux of TASEP-H

We now define the flux of TASEP-H, with or without perturbations. (TASEP-H without perturbations can be viewed as a special case of TASEP-H with perturbations, with $\lambda = 0$.)

Denote the state of the system at time t by $\overline{Z}(t) = (Z_0(t), \ldots, Z_{n-1}(t))$, where $Z_i(t) = 1$ if site i is a particle and $Z_i(t) = 0$ if it is a hole. Then, the instantaneous average flux of the system is defined as $\phi(\rho, n; \overline{Z}(t)) = \psi(\rho, n; \overline{Z}(t))/n$, where $\psi(\rho, n; \overline{Z}(t))$ is the expected total distance that will be traveled by all particles at time $t+1$. (This includes particles being relocated due to perturbations — the precise convention will be given below.) The *flux* of TASEP-H, with perturbations ($\lambda > 0$) or without perturbations ($\lambda = 0$), with parameters ρ and n, is defined as

$$\phi(\rho, n) \doteq \mathbb{E}\phi(\rho, n; \overline{Z}(\infty)),$$

where $\overline{Z}(\infty)$ is the (random) value of $\overline{Z}(t)$ in a stationary regime. Recall that in all cases, except $\lambda = 0$ (no perturbations) and $\rho \leq 1/2$, the stationary distribution is unique. In the remaining case of no perturbations and $\rho \leq 1/2$, recall that an ideal state is eventually reached, after which time all particle are free and move and speed 1; so, even though there may be multiple stationary distributions (corresponding to different ideal states), for each of them $\phi(\rho, n) = \rho$. Thus, in all cases, $\phi(\rho, n)$ above is well defined.

2.4. Asymptotic regime. Typical flux and formal flux

The asymptotic regime that we will consider is such that ρ and λ remain constant, while $n \to \infty$. To avoid cumbersome notation, let us assume that ρn is integer. This is the total number of particles. (If ρn is non-integer, we could assume that the number of particles is, say, $\lfloor \rho n \rfloor$.)

The main focus of the paper is on identifying the *limiting flux of TASEP-H with perturbations* ($\lambda > 0$):

$$\phi(\rho) = \lim_{n \to \infty} \phi(\rho, n).$$

The value of $\phi(\rho)$ will also be referred to as the *typical flux* of TASEP-H.

Formal flux of TASEP-H is defined as $\lim_{n \to \infty} \phi(\rho, n)$ for TASEP-H without perturbations ($\lambda = 0$). If $\rho < 1/2$, the formal flux is equal to ρ, simply because $\phi(\rho, n) = \rho$ for all n.

2.5. Process definition. Hydrodynamic scaling

For a system with parameter n (the circle length), the process state at time 0 is described by the function

$$F^n(x, 0) = \sum_{i=0}^{x} Z_i(0), \quad x = 0, 1, \ldots, n-1,$$

Clearly, $F^n(x, 0)$ is the total number of particles in sites $0, \ldots, x$ at time 0.

The (random) movement of particles in the system is described by the flux-function $\Phi^n(x, t)$, defined for $x = 0, 1, \ldots, n-1$ and time $t = 0, 1, 2, \ldots$ as follows: $\Phi^n(x, t)$ is the total number of particles that moved (right) from site x at times $1, \ldots, t$, with $\Phi^n(x, 0) = 0$. By convention, this quantity includes the number of particles that crossed site x due to perturbations, where, by convention, a perturbation moves a particle in a clock-wise direction. For example, if $n = 50$ and at time t particles moved from site 2 to 3 and from site 8 to 9 due to normal particle movement, and then a perturbation relocated a particle from site 48 to site 6, then the number of particles that moved from site 2 is two, while the number of particles that moved from

site 8 is one. This is because, by our convention, the particle relocated by the perturbation, simultaneously "leaves" sites 48, 49, 0, 1, 2, 3, 4, 5, in the clock-wise direction.[1]

We define

$$F^n(x, t) = F^n(x, 0) - \Phi^n(x, t), \quad x = 0, 1, \ldots, n - 1, \quad t = 0, 1, 2, \ldots$$

It is convenient to extend the definition of $F^n(x, t)$ to all integer x by setting $F^n(n, t) = F^n(0, t) + \rho n$ and assuming that $F^n(x, t)$ has periodic increments in x with period n. Clearly, the (random) function $F^n(x, t)$ completely describes the process evolution. It is non-decreasing in x and non-increasing in t. (In terms of describing the system states and evolution, only increments of $F^n(x, t)$ matter. So, for any fixed integer C, $F^n(\cdot, t) + C$ describes exactly the same system state at t as $F^n(\cdot, t)$, and $F^n(\cdot, \cdot) + C$ describes exactly the same system trajectory as $F^n(\cdot, \cdot)$. In particular, if $F^n(\cdot, t_1) - F^n(\cdot, t_2) = C$, $\forall x$, the system states at times t_1 and t_2 are equal.)

It is also convenient to extend the definition of $F^n(x, t)$ to all real $x \in \mathbb{R}$ and $t \geq 0$ by adopting convention

$$F^n(x, t) = F^n(\lfloor x \rfloor, \lfloor t \rfloor).$$

Furthermore, we will identify any location $x \in \mathbb{R}$ with $x \pmod{n} \in [0, n)$. So, for example, if $n = 50$, $F^{50}(4, t) - F^{50}(4 - 7, t)$ is the total number of particles, at time t, in the 7 consecutive sites, starting from 4 and going left: 4, 3, 2, 1, 0, −1, −2 or, equivalently, 4, 3, 2, 1, 0, 49, 48.

We assume that the system processes, for a given n, are constructed via certain driving (control) sequences of random variables. These driving sequences determine: the times it takes for particles at the right end of clusters to break away from their cluster; the timing of the perturbations; the choice of the particle to be relocated by each perturbation, and the choice of a vacant site where it relocates.

Before we specify the driving sequences, we will adopt the following convention about indexing (or, numbering) of clusters as the process evolves. For the clusters that exist initially (at time 0) we assign indices $j = 1, \ldots, k(0)$, where $k(0)$ is the number of initial clusters. When a new cluster is created due to a perturbed particle landing next to a free particle (or two free particles, on its both sides), this cluster is assigned the next available new index; specifically, if K is the maximum index assigned so far, the new cluster gets index $K + 1$ (and then K is reset to $K + 1$). Similarly, if a cluster j breaks into two, due to a perturbation, then, by convention, the right one retains the index j, and the left one is assigned the next available new index. If two clusters j' and j, separated by a single hole between them and such that j is located to the right of j', merge into one larger cluster due to a perturbation filling the hole separating them, then the combined cluster keeps the index j, while the index j' is discarded. If cluster j disappears (due to a normal evolution or perturbation), index j is also discarded. Thus, indices are assigned to new formed clusters in the increasing order, and the indices of disappearing clusters are discarded and never "reused". The description of driving sequences and their usage is as follows.

For each (cluster) index $j = 1, 2, \ldots$ we have an independent sequence of i.i.d. random variables, ξ_i^j, $i = 1, 2, \ldots$, with distribution $GEOM(1, p)$, mean $1/p$. R.v. ξ_i^j determines the random time it takes for the ith (since initial time 0) particle to leave the right end of cluster j, after it becomes the right-end particle. (This time has distribution $GEOM(1, p)$, as we already

---

[1] Note that the contribution of perturbations into the steady-state flux is upper bounded by $n(\lambda/n)/n$, and therefore vanishes as $n \to \infty$.

discussed in Section 1.3.) R.v. ξ_i^j, for $i \geq 2$, will also be the number of holes between ith and $(i-1)$-th leaving particles, if the $(i-1)$-th particle remains free at the time when the ith particle leaves the cluster.

The choice of the particle to be relocated by the ith perturbation (since initial time 0), $i = 1, 2, \ldots$, and the choice of a vacant site where the particle relocates, are determined by the r.v. $\eta_{1,i}$ and $\eta_{2,i}$, respectively; all these r.v. are i.i.d., each having uniform distribution in $[0, 1)$. Specifically, $\eta_{1,i}$ is "used" as follows. When ith perturbation needs to be applied, we label the particles by $1, 2, \ldots, \rho n$, in the order of their locations in the (scaled) interval $[0, 1)$. Then, the particle with label $\lceil \eta_{1,i} \rho n \rceil$ is picked to be relocated by the perturbation. The variable $\eta_{2,i}$ is used analogously for picking a vacant site where a particle relocates, except we label the currently vacant sites by $1, 2, \ldots, (1-\rho)n+1$, and pick the site with label $\lceil \eta_{2,i}[(1-\rho)n+1] \rceil$.

The timing of perturbations is determined by an i.i.d. sequence $\eta_{3,i}$, $i = 1, 2, \ldots$, with $\eta_{3,i} \sim GEOM(0, \lambda/n)$, $\mathbb{E}\eta_{3,i} = (1 - \lambda/n)/(\lambda/n)$. Namely, $\eta_{3,i}$ determines the random time from the start of the ith perturbation "clock", until the ith perturbation is activated. (This time has distribution $GEOM(0, \lambda/n)$, because, after the ith perturbation "clock" starts, the perturbation is activated in each slot with probability λ/n.) For I-perturbations, the ith perturbation clock starts in the time slot immediately after the previous, $(i-1)$-th, perturbation. For A-perturbations, the ith perturbation clock starts in the first time slot after the previous, $(i-1)$-th, perturbation, in which the process enters an ideal state.

For each n, we consider a hydrodynamic scaled process, where we compress space, time and the number of particles by factor n:

$$f^n(x, t) = \frac{1}{n} F^n(nx, nt).$$

For a fixed (scaled) time t, $f^n(x, t)$ – the (scaled) system state – has periodic increments with period 1. From now on, we always refer to scaled space and time, unless explicitly stated otherwise.

In this paper we study the asymptotic behavior of the hydrodynamic-scaled process $f^n(x, t)$ as $n \to \infty$.

In the rest of the paper, we will often construct the processes for all n on a common probability space. The construction is as follows. We assume that the i.i.d. sequences, ξ_i^j, $i = 1, 2, \ldots$, $j = 1, 2, \ldots$, are common for all n. They satisfy the following functional strong law of large numbers (FSLLN) condition: w.p.1,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \xi_i^j = \frac{1}{p} t, \quad t \geq 0 \ (u.o.c.), \quad \forall j. \tag{1}$$

The i.i.d. sequences $\eta_{1,i}$ and $\eta_{2,i}$, $i = 1, 2, \ldots$, will also be common for all n. The i.i.d. sequence $\eta_{3,i}^n$, $i = 1, 2, \ldots$, will depend on n. Using Skorohod representation, we can and will assume that these sequences are such that, w.p.1,

$$\lim_{n \to \infty} n^{-1} \eta_{3,i}^n = \hat{\eta}_{3,i}, \quad \forall i, \tag{2}$$

where $\hat{\eta}_{3,i}$ are independent (across i) exponentially distributed random variables with mean $1/\lambda$.

3. A-perturbations model

3.1. A-perturbations model: Hydrodynamic limits

For any n, $f^n(\cdot, t)$ is a Markov process, and the main goal of this paper is to study its asymptotic behavior, in particular the asymptotics of its stationary distribution. However, to

achieve this goal, we will need to analyze an extended version of this process, which includes additional process variables, explicitly describing clusters, sparse intervals, and the times when clusters exist. Recall that, for the A-perturbations model, *as far as the stationary distribution of the process is concerned,* the states with more than one cluster can be ignored; see Lemma 1 and the paragraph following it. Our main results concern the flux, limiting flux (typical flux) and formal flux, which are all defined in terms of stationary distributions. Therefore, for the A-perturbations model, it suffices to consider the process living on the states with at most one cluster, and so the following additional descriptors will suffice.

Denote by $\ell^n(t)$ and $r^n(t)$ the (scaled) locations of the particles at the left and right edge of the cluster at time t, if any, and by $\tau^n(t) = r^n(t) - \ell^n(t)$ its length. (Note that the scaled number of particles within the cluster is $[n\tau^n(t) + 1]/n = \tau^n(t) + 1/n$, not $\tau^n(t)$.) By convention, if cluster does not exist at time 0, we assume that $\ell^n(t) = r^n(t) = \tau^n(t) = 0$ until the time when a cluster forms. Also by convention, after a cluster dissolves, $\ell^n(t) = r^n(t)$ remain frozen at the value of ℓ^n at the last time when the cluster existed. We also adopt the convention that when ℓ^n changes as a result of a cluster formation, the new value of ℓ^n is chosen to be within $[0, 1)$, and r^n is chosen accordingly, to be to the right of (or at) ℓ^n.

At any time t, the system state is given by

$$[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)].$$

(Note that $[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)]$ contains no more information than $f^n(\cdot, t)$, in the sense that the distribution of $f^n(\cdot, s), s \geq t$ depends only on $f^n(\cdot, t)$. Once again, we use the additional state descriptors for the purposes of analysis.) The metrics on the state space components are as follows. For the f-component, it is defined by the max-norm:

$$\|f_1(\cdot) - f_2(\cdot)\| = \max_x |f_1(x) - f_2(x)|.$$

For the τ-component, it is given simply by the distance $|\tau_1 - \tau_2|$. For the ℓ-component, it is the distance between $\lfloor \ell_1 \rfloor$ and $\lfloor \ell_2 \rfloor$, considered as points on the unit length circle; formally, it is

$$\min_{k \in \mathbb{Z}} |\ell_1 - \ell_2 + k|.$$

Finally, the metric on the entire system state space is the sum-metric for its f-, τ- and ℓ-components. (There is no need to consider the metric for the r-component.)

For each n, consider a finite or infinite time interval $[0, T^n]$, with T^n possibly depending on n. (If T^n does depend on n, it can be thought of as a realization of a stopping time.) Trajectories

$$[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)], \quad t \in [0, T^n],$$

are considered as elements of the Skorohod space. For the purposes of having the Skorohod metric well-defined, we adopt the convention that, when $T^n < \infty$, the state remains constant in $[T^n, \infty)$.

Further, let $\alpha_0^n \geq 0$ be the first time when a cluster exists. (If it exists at time 0, then $\alpha_0^n = 0$. If it never exists $\alpha_0^n = \infty$.) Let $\beta_0^n \geq \alpha_0^n$ be the first time after α_0^n when the system is in an ideal state. ($\beta_0^n = \infty$ if the system never enters an ideal state after α_0^n.) In other words, $[\alpha_0^n, \beta_0^n)$ is the time interval when the "first" (in time) cluster exists. Similarly, let $\alpha_1^n \geq \beta_0^n$ be the first time after β_0^n when the "second" cluster appears, and $\beta_1^n \geq \alpha_1^n$ be the first time this cluster disappears. And so on, we define pairs (α_i^n, β_i^n) marking the beginning and end of the ith cluster. Note that, $\alpha_i^n = \infty$ implies that all consecutive times α_j^n and β_j^n are also infinite. We also adopt a convention that, if we consider the process on a finite time interval $[0, T^n]$,

then any time α_i^n or β_i^n which is outside $[0, T^n]$ is set to be infinite. Finally, note that all α_i^n, except maybe α_0^n when it is 0, are the times when perturbations occur. (However, not every perturbation time is one of the α_i^n.)

The following will be called an extended realization (trajectory) of the process on a (finite or infinite) time interval $[0, T^n]$:

$$\{[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)], \ t \in [0, T^n]; \quad [\alpha_i^n, \beta_i^n], i = 0, 1, \ldots\}.$$

(When this does not cause confusion, which is in most cases, we will call it just a realization.)

Definition 1 (*Hydrodynamic Sample Path (HSP)*)**.** Suppose, there is a fixed sequence of the process (extended) realizations on the time intervals $[0, T^n]$

$$\{[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)], \ t \in [0, T^n]; \quad [\alpha_i^n, \beta_i^n], i = 0, 1, \ldots\}, \tag{3}$$

such that the driving sequences' realizations satisfy conditions (1)–(2). Then, a trajectory

$$\{[f(\cdot, t), \ell(t), r(t), \tau(t)], \ t \in [0, T]; \quad [\alpha_i, \beta_i], i = 0, 1, \ldots\}, \tag{4}$$

is called a *hydrodynamic sample path* (HSP) on the interval $[0, T]$ if the following conditions hold.

(i) Points $0 \le \alpha_0 \le \alpha_1 \le \ldots$ are such that there is only a finite number of them on any finite interval; $\alpha_i < \alpha_{i+1}$, as long as $\alpha_i < \infty$; if $\alpha_i > T$, then necessarily $\alpha_i = \infty$. Points $0 \le \beta_0 \le \beta_1 \le \ldots$ are such that there is only a finite number of them on any finite interval; $\alpha_i \le \beta_i < \alpha_{i+1}$, as long as $\beta_i < \infty$; if $\beta_i > T$, then necessarily $\beta_i = \infty$.

(ii) Function $f(x, t)$ is jointly continuous in (x, t). Function $\tau(t)$ is continuous in t. Function $[\ell(t), r(t)]$ is RCLL in t; its only possible points of discontinuity are those α_i, $i \ge 1$, that are finite.

(iii) Trajectory (4) is a limit of trajectories (3), as $n \to \infty$, in the following sense:

(iii.1) $T^n \to T$;

(iii.2) $\alpha_i^n \to \alpha_i$ and $\beta_i^n \to \beta_i$, for each i;

(iii.3) The following u.o.c. convergence

$$\{[f^n(x, t), \tau^n(t)], \ x \in \mathbb{R}, \ t \ge 0\} \to \{[f(x, t), \tau(t)], \ x \in \mathbb{R}, \ t \ge 0\} \tag{5}$$

and the following Skorohod space convergence

$$\{[\ell^n(t), r^n(t)], \ t \ge 0\} \to \{[\ell(t), r(t)], \ t \ge 0\} \tag{6}$$

hold. (In (5) and (6), by convention, $[f(\cdot, t), \ell(t), r(t), \tau(t)]$ is defined for all $t \ge 0$, with its value being constant for $t \ge T$, and similarly $[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)]$ is constant for $t \ge T^n$.)

For future reference, note the following. Given that α_i are the only possible points of discontinuity of $[\ell(t), r(t)]$, the Skorohod convergence (6) implies that the convergence

$$[\ell^n(t), r^n(t)] \to [\ell(t), r(t)] \tag{7}$$

is uniform on any closed bounded interval, not containing any of the points α_i.

The following lemma describes basic properties of the HSPs, implied by the corresponding basic properties of pre-limit trajectories and HSP definition.

Lemma 2. *Any HSP has the following properties.*

(i) Function $f(x, t)$ is 1-Lipschitz non-increasing in t and 1-Lipschitz non-decreasing in x.

(ii) $f(x + 1, t) - f(x, t) = \rho$ for any x and t.

(iii) Function $\tau(t)$ is 1-Lipschitz.

(iv) Functions $\ell(t)$ and $r(t)$ are 1-Lipschitz non-increasing on the intervals of the form $[\alpha_i, \beta_i \wedge T)$. In sub-intervals of $[0, T)$, not intersecting with any of the intervals $[\alpha_i, \beta_i \wedge T)$, $\tau(t) = 0$ and $\ell'(t) = r'(t) = 0$.

(v) For any $t > 0$, condition $\tau(t) > 0$ necessarily implies that $t \in (\alpha_i, \beta_i \wedge T]$ for some i.

(vi) For any t, $f(r(t), t) - f(\ell(t), t) = r(t) - \ell(t) = \tau(t)$.

(vii) For any t, $f(x, t)$ is $1/2$-Lipschitz in $[r(t), \ell(t) + 1]$.

Proof is rather straightforward. Let us prove (i). Note that for any n, $\delta \geq 0$, x, and $t + \delta < T^n$,

$$0 \leq f^n(x + \delta, t) - f^n(x, t) \leq \delta + 2/n, \tag{8}$$

$$-\delta - 2/n - \kappa/n \leq f^n(x, t + \delta) - f^n(x, t) \leq 0, \tag{9}$$

where κ is the (finite) number of those α_i which are in $[0, t + \delta]$. Taking $n \to \infty$ limit, and using the u.o.c. convergence (5) and monotonicity of $f^n(x, t)$, we obtain the monotonicity and Lipschitz properties of $f(x, t)$. This proves (i).

Proof of (ii) follows from the fact that $f^n(x + 1, t) - f^n(x, t) = \rho$ holds for any n.

To prove (iii) it suffices to note that $\tau^n(t)$ can change at most by $3/n$ at times α_i^n and β_i^n; at other times it can change at most by $1/n$.

The remaining statements (iv)–(vii) are proved analogously, using in particular the following obvious properties of pre-limit trajectories:

$r^n(t)$ and $\ell^n(t)$ can only decrease by $1/n$ at any time t, except the times α_i^n and β_i^n; for any t,

$$\tau^n(t) = r^n(t) - \ell^n(t) + O(1/n),$$
$$f^n(r^n(t), t) - f^n(\ell^n(t), t) = r^n(t) - \ell^n(t) + O(1/n),$$

and

$$f^n(x_2, t) - f^n(x_1, t) = (x_2 - x_1)/2 + O(1/n), \quad \forall x_1, x_2 \text{ such that } r(t) < x_1 < x_2 < \ell(t) + 1.$$

Taking $n \to \infty$ limits, and using in particular (7), we easily obtain (iv)–(vii). We omit further details. $\square$

For an HSP, a time point t is called *regular*, if the derivatives $\tau'(t)$, $\ell'(t)$, $r'(t)$ exist. In view of Lemma 2(iv)–(v), *almost all points (w.r.t. Lebesgue measure) are regular.*

We see that an HSP describes the evolution of the distribution of the continuous "particle mass", or "fluid", obtained as a limit of rescaled pre-limit particle distributions. For an HSP we will also use the natural notion of a cluster: we say that a cluster $[\ell(t), r(t)]$, of length $\tau(t) = r(t) - \ell(t)$, exists at time t, if $t \in [\alpha_i, \beta_i)$ for some i; in this case the rest of the unit circle, that is interval $[r(t), \ell(t) + 1]$ is a sparse interval. A sub-interval of a sparse interval we will call a sparse sub-interval. Note that it is possible that a cluster in an HSP exists even when it has zero length, $\tau(t) = 0$. Of course, if $\tau(t) > 0$ then a cluster necessarily exists. If there is no cluster at time t, then necessarily $\tau(t) = 0$ and the entire unit circle $[0, 1]$ is a sparse interval. The total particle mass $f(x + 1, t) - f(x, t) = \rho$ is constant at all times. The particle mass density $(\partial/\partial x)f(x, t)$ within a cluster is exactly 1, and the density within a sparse interval is at most $1/2$.

Lemma 3 describes the basic HSP dynamics. We precede it with an informal description of this dynamics. (See Fig. 3 for an illustration.) The particle mass outside a cluster (if any) simply moves to the right at the constant speed 1 until and unless it "hits" the cluster. In a time interval when a cluster $[\ell(t), r(t)]$ exists, its right end $r(t)$ moves left at the constant speed p.
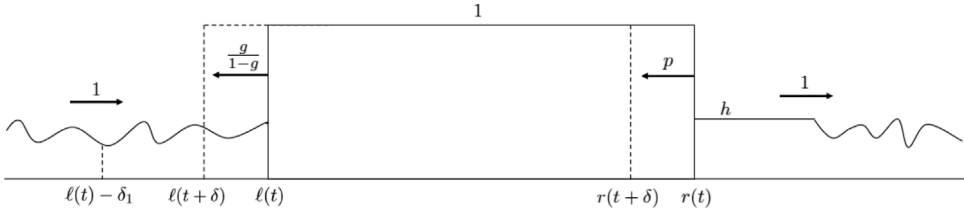
Fig. 3. Cluster dynamics illustration.

(Because this is the rate at which particles in a pre-limit system "break-away" from the cluster on the right.) Moreover, as the right end of a cluster moves left, it leaves a sparse sub-interval of the density exactly $h = p/(1 + p)$ in its "wake". (Because, in pre-limit process, the average distance between two consecutive breaking away particles is $1 + 1/p = 1/h$; see Sections 1.3 and 2.5.) Specifically, in a time interval $[t, t + \delta]$, a sparse sub-interval $[r(t) - \delta p, r(t) + \delta]$ of density h is created, where $r(t + \delta) = r(t) - \delta p$ is the right-end location at time $t + \delta$, and $r(t) + \delta$ is how far the right edge of the sparse sub-interval moved right by time $t + \delta$. Now let us consider how the left end $\ell(t)$ moves. Consider a small interval $[t, t + \delta]$. Let $[\ell(t) - \delta_1, \ell(t)]$ be the sparse sub-interval immediately to the left of the cluster at time t, which hits – and merges into – the cluster in time interval $[t, t + \delta]$. Then the δ_1 above is the unique solution of equation

$$\delta = \delta_1 - [f(\ell(t), t) - f(\ell(t) - \delta_1, t)]. \tag{10}$$

Indeed, the time δ for the sparse sub-interval $[\ell(t) - \delta_1, \ell(t)]$ (at t) to join the cluster is exactly equal to the "empty space" in $[\ell(t) - \delta_1, \ell(t)]$ (at t), which is $\delta_1 - [f(\ell(t), t) - f(\ell(t) - \delta_1, t)]$; solution δ_1 is unique, because the RHS of (10), as function of δ_1, is Lipschitz increasing with derivative $\geq 1/2$ (recall that the density within a sparse interval is at most $1/2$).

The following lemma describes the above basic dynamics of an HSP formally. Its proof is also fairly straightforward — most of it uses the corresponding properties of pre-limit trajectories and HSP definition. We will only provide the key points and comments, omitting details.

Lemma 3. *Any HSP has the following properties.*
(i) Consider a time point t such that $t \in (\alpha_i, \beta_i \wedge T)$. (Recall that for a $0 < t < T$, $\tau(t) > 0$ necessarily implies $t \in (\alpha_i, \beta_i \wedge T)$.) For any sufficiently small time $\delta > 0$, the following holds. Denote by δ_1 the unique positive solution to equation

$$\delta = \delta_1 - [f(\ell(t), t) - f(\ell(t) - \delta_1, t)]. \tag{11}$$

Then, the HSP state at time $t + \delta$ is as follows:

$$\ell(t + \delta) = \ell(t) - (\delta_1 - \delta); \tag{12}$$
$$r(t + \delta) = r(t) - p\delta; \tag{13}$$

for $x \notin (\ell(t + \delta), r(t + \delta))$ (this is the part of sparse interval that just shifted, without any interaction with the cluster) we have

$$f(x, t + \delta) = f(x - \delta, t);$$

for $x \in (\ell(t + \delta), r(t + \delta))$ *(within cluster at time* $t + \delta$*)*

$$(\partial / \partial x) f(x, t + \delta) = 1;$$

for $x \in (r(t + \delta), r(t) + \delta)$ *(the "wake" of right end of cluster)*

$$(\partial / \partial x) f(x, t + \delta) = h.$$

In addition, if this t *is regular,*

$$\ell'(t) = -\frac{g}{1 - g}, \quad \text{where } g = g(t) = \frac{\partial^-}{\partial x} f(\ell(t), t) \tag{14}$$

and $r'(t) = -p$.

(ii) *If* t *is regular, then* $\tau'(t) > 0$ *if and only if* $g > h$.

(iii) *If* t *is regular,* $t \in (\alpha_i, \beta_i \wedge T)$ *and* $\tau(t) = 0$*, then, necessarily,* $\tau'(t) = 0$ *and* $g = h$.

(iv) *For any interval* $[t, t + \delta] \subset [0, T)$ *when a cluster does not exist (i.e. not intersecting with any of the intervals* $[\alpha_i, \beta_i)$*), the particle density moves to the right at speed* 1:

$$f(x, t + \delta) = f(x - \delta, t), \quad \ell(t + \delta) = r(t + \delta) = \ell(t) + \delta = r(t) + \delta, \quad \tau(t) = 0.$$

(v) *Denote by* $\mu(t)$ *the total Lebesgue measure of all points* $x \in [0, 1)$ *(within the circle) at time* t*, which are outside the cluster* $[\ell(t), r(t)]$ *(if any) and where the density*

$$\frac{\partial}{\partial x} f(x, t) > h.$$

Function $\mu(t)$ *is Lipschitz (and then the derivative exists for almost all* t*), and given the properties* (i)–(iv)*, it is non-increasing. Then, for almost all* $t > 0$ *(w.r.t. Lebesgue measure), either*

$$\mu'(t) = 0,$$

or this time t *is regular,* $\tau'(t) > 0$ *(and then* $\tau(t) > 0$*,* $t \in (\alpha_i, \beta_i \wedge T)$*), and*

$$g > h \quad \text{and} \quad \mu'(t) = -1 + \ell'(t) = -\frac{1}{1 - g} \leq -\frac{1}{1 - h} = -(1 + p).$$

(vi) *The Lebesgue measure of the time points* t *where* $\tau'(t) > 0$ *is upper bounded by* $1/(1 + p)$.

From now on we make the definition of a regular point more restrictive by requiring, in addition, that the conclusions of Lemma 3(v) hold. Still, almost all time points t are regular.

Proof. (i) The meaning of (11) (which is same as (10)) is explained in the informal definition above. Consider a fixed $\delta > 0$ and time t. For each pre-limit trajectory, as in HSP definition, define $\delta_1^{(n)} \geq 0$ as the smallest number such that all particles located at time t in the interval $[\ell^n(t) - \delta_1^{(n)}, \ell^n(t)]$, will join the cluster by time $t + \delta$. It is easy to observe that

$$\delta = \delta_1^{(n)} - [f(\ell^n(t), t) - f(\ell^n(t) - \delta_1^{(n)}, t)] + O(1/n). \tag{15}$$

Indeed, up to quantities that are $O(1/n)$, the $\delta_1^{(n)}$-long (scaled) interval $[\ell^n(t) - \delta_1^{(n)}, \ell^n(t)]$ at time t consists of (scaled) number $[f(\ell^n(t), t) - f(\ell^n(t) - \delta_1^{(n)}, t)]$ of particles and (scaled) number δ of holes, because the rate at which holes "eliminated" by the left end of the cluster is exactly 1. (In the unscaled trajectory, exactly one hole is eliminated per time slot; in the scaled trajectory, $1/n$ of empty space is eliminated per $1/n$-long time slot; so the rate of empty space

elimination is 1 for both the unscaled and scaled trajectories.) It remains to take the $n \to \infty$ limit in (15) (as in the HSP definition) and observe that $\delta_1^{(n)}$ must converge exactly to the δ_1 as defined, thus proving (11). All properties before (14) then easily follow. Property (14) is a differential form of (11). Namely, we use the fact that t is regular and therefore $\ell'(t)$ exists; let $\delta \downarrow 0$; from (12) we have

$$\lim_{\delta \to 0} \delta_1/\delta = 1 - \ell'(t);$$

dividing (11) by δ, taking $\delta \downarrow 0$ limit and substituting the above display, we obtain (14). The last property is from (13).

(ii) and (iii) follow from (i). (iv) is obvious.

(v) Obviously, $\mu'(t) = 0$ for almost all time points t such that a cluster does not exist. Consider now a t which is strictly within a time interval when a cluster exists. Once again, consider the movement of the left end ℓ of the cluster in a small interval $[t, t+\delta]$. Then, using (14), (11) can be written as

$$- \delta_1 = \int_t^{t+\delta} [-1 + \ell'(s)]ds = \int_t^{t+\delta} \frac{-1}{1 - g(s)}ds, \tag{16}$$

where $-1 + \ell'(s)$ is the instantaneous (negative) rate at which the sparse sub-interval $[\ell(t) - \delta, \ell(t)]$ of length δ_1, which existed at time t, "shrinks" due to its right end $\ell(s)$ moving at speed $\ell'(s)$ left and its left end $\ell(t) - \delta_1 + (s - t)$ moving at speed 1 right. The integrand in the RHS exists for almost all t (and is unique up the time subsets of zero Lebesgue measure). Note that, for a given t, the density $(\partial/\partial x)f(x, t)$ exists a.e. in x, and therefore $(\partial^-/\partial x)f(x, t) = (\partial/\partial x)f(x, t)$ a.e. (Recall that $g(t)$ is defined as left derivative.) Analogously to (16), it is easy to obtain its generalization, which gives the (negative) increment of μ in the interval $[t, t+\delta]$, by integrating over only those times s, where $(\partial^-/\partial x)f(\ell(s), s) = g(s) > h$:

$$\mu(t + \delta) - \mu(t) = \int_t^{t+\delta} \frac{-1}{1 - g(s)} \mathbb{I}\{g(s) > h\}ds. \tag{17}$$

Therefore, for almost all $s \in [t, t + \delta]$,

$$\mu'(s) = \frac{-1}{1 - g(s)} \mathbb{I}\{g(s) > h\}.$$

From here, along the fact that almost all times s are regular, all properties stated in (v) follow.

(vi) Follows from (v). $\quad \square$

Lemma 4. *Suppose, $\rho < h$. Then any HSP is such that the following properties hold.*
(i) For any i such that $\alpha_i < \infty$, we have $\beta_i \wedge T - \alpha_i < 1$.
Consequently, for any $t_0 \leq T - 1$ there exist $t_0 < t_1 \leq t_0 + 1$ such that $\tau(t_1) = 0$.
(ii)

$$\int_0^T \tau(t)dt \leq \bar{C} = 1 + \frac{1}{2(1 + p)}. \tag{18}$$

Proof. (i) Suppose not, i.e. $(t_0, t_0 + 1]$ is entirely within an interval of the form $[\alpha_i, \beta_i \wedge T)$. Then, at time $t_0 + 1$ the HSP state is such that everywhere outside the cluster the density is equal to h. (This follows from Lemma 3(i), density h.) This is impossible, because it would imply that $\rho \geq h$.

(ii) Consider the set of points $t \in [0, T)$, where $\tau(t) > 0$. This set consists of possibly an interval $[0, s)$, whose length s cannot exceed 1 (by (i)), and a countable number of open intervals (t_1, t_2) with lengths also not exceeding 1 (by (i)). Obviously,

$$\int_0^s \tau(t)dt \le s \le 1.$$

Consider any of the open intervals (t_1, t_2). We have

$$\max_{t \in (t_1, t_2)} \tau(t) \le \int_{t \in (t_1, t_2): \ \tau'(t) > 0} \tau'(t)dt \le \int_{t \in (t_1, t_2): \ \tau'(t) > 0} [-\ell'(t)]dt$$
$$\le \frac{1}{2}\mathcal{L}\{t \in (t_1, t_2): \ \tau'(t) > 0\},$$

and therefore (recall that $t_2 - t_1 \le 1$)

$$\int_{t_1}^{t_2} \tau(t)dt \le \frac{1}{2}\mathcal{L}\{t \in (t_1, t_2): \ \tau'(t) > 0\}.$$

Summing up over all intervals where $\tau(t) > 0$, we obtain

$$\int_0^T \tau(t)dt \le 1 + \frac{1}{2}\mathcal{L}\{t \in (0, T): \ \tau'(t) > 0\} \le 1 + \frac{1}{2}\frac{1}{1 + p},$$

where in the last inequality we used Lemma 3(vi). □

Suppose $h < \rho < 1$. Recall $h = p/(1 + p)$. State $[f^*(\cdot), \ell^*, r^*, \tau^*]$ is called a *main equilibrium state* (MES), if it satisfies the following conditions: $\tau^* = \frac{\rho - h}{1 - h}$, $r^* = \ell^* + \tau^*$, $f'(x) = 1$ for $x \in (\ell^*, r^*)$, $f'(x) = h$ for $x \in (r^*, \ell^* + 1)$. The definition shows that a MES is essentially unique, up to a space shift of the cluster; therefore, we will often refer to any MES as *the* MES. Note that if the steady-state of a system with large n is close to MES, then the flux is close to $\phi^* = \phi^*(\rho) \doteq (1 - \tau^*)h = p(1 - \rho)$.

Lemma 5. *Suppose, $\rho > h$. Then any HSP is such that the following properties hold.*

(i) If HSP initial state is a MES, the HSP is unique and is stationary, i.e. it stays in this state (up to shift of the cluster).

(ii) If $\alpha_0 = 0$ and $\beta_0 \ge 1 - \rho$, then the HSP state at time $1 - \rho$ is a MES, and the trajectory is stationary (staying in a MES) in the interval $[1 - \rho, T]$.

Proof. (i) According to derivatives' expressions for a MES, it cannot change with time, up to the cluster moving left at the constant speed p.

(ii) The entire particle mass, which is originally outside the cluster is at most ρ. We have $\beta_0 \ge 1 - \rho$, so the cluster exists for the time at least $1 - \rho$. The time interval $[0, 1 - \rho]$ is long enough for all the mass which was originally outside the cluster to join the cluster. Therefore, at some time point in $t_0 \in [0, 1 - \rho]$, a state is reached, consisting of a cluster and a sparse interval with constant density h (by Lemma 3(i)) This is equivalent to the state being a MES. The same argument as in the proof of (i) shows that the trajectory must stay in MES in $[t_0, T]$. □

Lemma 6. *Consider a sequence of systems, with $n \to \infty$, with arbitrary random initial states*

$$[f^n(\cdot, 0), \ell^n(0), r^n(0), \tau^n(0)].$$

For each n consider some stopping time T^n (finite or infinite). Then, with prob. 1, any subsequence of (extended) realizations has a further subsequence, along which the convergence to an HSP holds in the sense of Definition 1.

Proof. This type of a hydrodynamic-limit result is standard. If the sequence of processes is constructed on the common probability space, as specified above, the driving sequences satisfy the properties required in Definition 1 w.p.1. Clearly, w.p.1 we can and do choose a subsequence along which the convergence in (iii.1) and (iii.2) of Definition 1 holds, where all finite α_i are different. Properties (8) and (9) hold for *any* pre-limit trajectory. Then, w.p.1 we can choose a further subsequence along which the convergence (5) in Definition 1 holds for all rational pairs (x, t); given the monotonicity of $f^n(x, t)$ (non-decreasing in x and non-increasing in t) we see that the convergence (5) is in fact u.o.c., as required in Definition 1. Using the properties of $r^n(t)$ and $\ell^n(t)$ described in the proof of Lemma 2, w.p.1 we can choose a further subsequence along which the Skorohod convergence (6) holds. Then, by Definition 1, the limiting trajectory is an HSP. □

Lemma 7. *Suppose $\rho > h$. Consider a sequence of systems, with $n \to \infty$, with arbitrary random initial states $[f^n(\cdot, 0), \ell^n(0), r^n(0), \tau^n(0)]$. For a fixed constant $B \geq 1 - \rho$ and each n consider a stopping time T^n, which is the minimum of constant B and the first time after 0, when the process is in ideal state. Then, w.p.1, any subsequence of (extended) realizations has a further subsequence, along which it converges to an HSP of duration $T \leq B$. This HSP satisfies one of the following two properties:*

(a) $T < 1 - \rho$ and $\tau(T) = 0$;
(b) $T = B$ and the HSP state in $[1 - \rho, B]$ is a MES.

Proof. The probability 1 subsequential convergence to HSPs is due to Lemma 6. Then the required HSP properties follow from Lemma 5(ii). □

3.2. A version of the ballot theorem

The following combinatorial fact plays a key role in the proofs of our main results for the high density, $h < \rho < 1/2$. We will derive it from Theorem 1.2.5 – a ballot theorem – in [11].

Proposition 8. *Suppose an integer $n \geq 1$ and a sequence of real non-negative numbers, $k_1, k_2, \ldots, k_n$ are fixed. Let us extend the definition of the sequence $\{k_r\}$ to all integer $r \geq 1$ by periodicity, $k_{r+n} = k_r$, and denote $\psi(j) = \sum_{\ell=1}^{j} k_\ell$ for $j \geq 0$, where $\psi(0) = 0$ by convention. Suppose a real number $m > \psi(n)$ is fixed. Let N be the number of those indices $j \in \{0, 1, \ldots, n - 1\}$, for which*

$$\psi(j + r) - \psi(j) < \frac{m}{n} r, \quad r = 1, \ldots, n. \tag{19}$$

Then,

$$N \geq \left\lceil n \left(1 - \frac{\psi(n)}{m} \right) \right\rceil. \tag{20}$$

The proof is in Appendix A.

3.3. A-perturbations model: Main results

Theorem 9.   *Consider the model with A-perturbations. Assume $\rho < h$. Then, as $n \to \infty$, the sequence of the stationary distributions of $[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)]$ is such that its any subsequential weak limit is concentrated on the states with $\tau = 0$. In other words, as $n \to \infty$, $\tau^n(\infty) \Rightarrow 0$ [equivalently, $\mathbb{E}\tau^n(\infty) \to 0$] and, consequently, the limiting flux (typical flux) $\phi(\rho) = \rho$.*

Proof of Theorem 9.   It suffices to show that any distributional limit is such that $\mathbb{E}\tau = 0$. Suppose not. Consider a fixed interval $[0, B]$, and stationary versions of the processes in this interval. Then there exists $\epsilon > 0$ such that, for any $B > 0$,

$$\liminf_{n \to \infty}(1/B)\mathbb{E}\int_0^B \tau^n(t)dt \geq \epsilon > 0. \tag{21}$$

However, using our construction above, we can construct all these (stationary) processes on a common probability space, so that, w.p.1., we have a subsequential convergence to an HSP. Convergence to an HSP, in particular, means that $\tau^n(\cdot) \to \tau(\cdot)$ u.o.c. Then, for any $B > 0$, from Lemma 4(ii) we obtain

$$\limsup_{n \to \infty}\mathbb{E}\int_0^B \tau^n(t)dt \leq \bar{C}.$$

This contradicts the fact that (21) must hold for all $B > 0$.   □

Lemma 10.   *Consider the model with A-perturbations. Assume $h < \rho < 1/2$. Consider the system without time/space rescaling. Consider a (random) initial system state, at initial time 0, which is formed as follows: we pick an arbitrary ideal state and apply a perturbation to it. (In other words, a particle is picked uniformly at random, and it is placed into one of the empty sites, chosen uniformly at random.) Then, there exists $\delta = \delta(p, \rho) > 0$ such that, uniformly in all the original (pre-perturbation) ideal states and uniformly in all sufficiently large n (and then in all n), the following event occurs with probability at least δ: The perturbation creates a cluster and this cluster will not disappear within (unscaled) time interval $[0, (1 - \rho)n - 1]$.*

Proof.   It will be convenient to relabel time, so that the initial time is 1 (instead of 0). Consider the initial state after the perturbation. Consider the relocated particle, which we will refer to as a "seed" particle; the site to which it relocates we will call the "seed" site. If the perturbation happens to form a cluster, let us consider the dynamics of this cluster, starting the initial time 1. (If the perturbation did not form a cluster, let us view such event as a cluster "disappearance" immediately at time 1.) Obviously, as long as the cluster exists, the particles will leave it on the right as a Bernoulli process, with "success" (leaving) probability p, starting time 2. (By convention, a particle does *not* leave the cluster at initial time 1.) Of course, the process of "failures" (non-departures) is Bernoulli with probability $1 - p$. Fix $\epsilon = (\hat{p} - p)/3$, where $\hat{p} = \rho/(1 - \rho) > p$, because $p = h/(1 - h)$. (The meaning of $\hat{p}$ will be explained shortly.) Let $D(t), \ t = 1, 2, \ldots$, be the cumulative number of "successes" of the Bernoulli process described above (with $D(1) = 0$ by convention), by and including time t, and let $\bar{D}(t) = t - D(t)$ be the corresponding cumulative number of "failures" (with $\bar{D}(1) = 1$ by convention). We observe that

$$\delta_1 = \mathbb{P}\{\bar{D}(t) \geq [(1 - p) - \epsilon]t, \ \forall t \geq 1\} > 0. \tag{22}$$

Now consider how the cluster (if any) "grows" on the left due to particles joining it on the left. (By convention, if the seed particle finds another one immediately to the left of the seed site, we assume that the latter particle joined the cluster immediately at the initial time 1.) Note that the initial cluster (if any) has at most 1 particle to the right of the seed particle and, therefore, at time 1, at least $\rho n - 3$ particles (i.e., "almost all") form a sparse interval to the left of the cluster. Therefore, at least $\rho n - 2$ particle which are initially to the left of the seed particle will move unobstructed, at speed 1, until and unless they join the cluster.

Let us define the following process over the time interval $1, 2, \ldots, n - (\rho n - 1) = (1 - \rho)n + 1$; this process exactly describes the process $A(t)$ of the cumulative number of particles joining the cluster from the left, in the (a bit shorter) time interval $1, 2, \ldots, (1 - \rho)n$. We will illustrate the definition by an example. Let $n = 20$ and the total number of particles is $\rho n = 8$. Consider the configuration of the $\rho n - 1 = 7$ particles on the circle, *excluding the seed particle*, with respect to the seed site. Starting from the site immediately to the left from the seed site, and moving left, the configuration of the particles (excluding seed particle) is, for example, this sequence of length $n = 20$:

$$1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0 \tag{23}$$

(Note that the last element of this sequence is necessarily 0, because it corresponds to the seed site, and we do *not* include the seed particle into this sequence.) Consider now the sequence of length $n - (\rho n - 1) = (1 - \rho)n + 1 = 20 - 7 = 13$, obtained from (23) by removing each 0, which immediately follows a 1. We obtain

$$1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 1, 1, 1 \tag{24}$$

The cumulative number of 1's by time t (= place in the sequence) is the function $A(t)$, $t = 1, 2, \ldots, (1 - \rho)n + 1$. Consider the interval $t = 1, 2, \ldots, (1 - \rho)n$ (which is a bit shorter than $t = 1, 2, \ldots, (1 - \rho)n + 1$). It is easy to see that, in this interval, until and unless the cluster disappears, $A(t)$ is exactly the number of particles joining the cluster on the left. This is because at the initial time 1 the particles to the left of the seed particle (except maybe the particle immediately to the left, which, by our convention, has joined the cluster at initial time 1) are part of the sparse interval and therefore will move at speed 1 until/unless they join the cluster. Using our illustration above, if the initial configuration of particles to the left of the seed particle is given by (23) then the particles will join the cluster at times $t = 1, 3, 7, 8, 11, 12$ (see the first $(1 - \rho)n = 12$ elements of sequence (24)), until/unless the cluster disappears at some $t \leq 12$. Here we consider the shorter time interval $t = 1, 2, \ldots, (1 - \rho)n$, as opposed to $t = 1, 2, \ldots, (1 - \rho)n + 1$, because this excludes from consideration the initially non-free particle (if any) immediately to the right of the seed particle.

Clearly, the process $A(t)$ satisfies the following conservation law: $A((1 - \rho)n + 1) = \rho n - 1$. Let us denote by $\bar{A}(t) = t - A(t)$ the corresponding cumulative process of "failures" (zeros). The conservation law for $\bar{A}(t)$ is: $\bar{A}((1 - \rho)n + 1) = ((1 - \rho)n + 1) - (\rho n - 1) = (1 - 2\rho)n + 2$. Note that the average slope of $\bar{A}(t)$ is

$$\frac{(1 - 2\rho)n + 2}{(1 - \rho)n + 1} = \frac{(1 - 2\rho) + 2/n}{(1 - \rho) + 1/n}, \tag{25}$$

and is close to

$$\frac{1 - 2\rho}{1 - \rho} = 1 - \hat{p}, \quad \hat{p} = \rho/(1 - \rho).$$

Therefore, for all large n, the average slope (25) is at most $(1 - \hat{p}) + \epsilon < (1 - \hat{p}) + 2\epsilon = (1 - p) - \epsilon$. Notice that, given locations of all particles except the seed particle on the circle, since the

seed particle chooses one of the empty sites (to become the seed site) uniformly at random, all cyclical permutations of a sequence (24) appear with equal probabilities. Using the latter observation, we can apply Proposition 8, to obtain the following estimate: for all sufficiently large n,

$$\mathbb{P}\{\bar{A}(t) < [(1 - \hat{p}) + 2\epsilon]t, \ \forall \ 1 \le t \le (1 - \rho)n + 1\} \ge \delta_2 = \frac{\epsilon}{(1 - \hat{p}) + 2\epsilon} > 0. \qquad (26)$$

Combining estimates (22) and (26), we conclude that, uniformly on all initial (pre-perturbation) ideal states and in all sufficiently large n, with probability at least $\delta = \delta_1 \delta_2$, a perturbation will create a cluster at time 1 and this cluster will not disappear by and including time $(1 - \rho)n$. $\quad\square$

As a corollary of Lemmas 7 and 10, we obtain Lemma 11. To formulate it, note the following. Recall the definition of the state space of $[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)]$ in Section 3.1, and the metric on it. For each state we can consider the unique "standard" version of each state, satisfying $f^n(0, t) = 0$, $\ell^n(t) \in [0, 1)$, and $r^n(t) \in [\ell^n(t), \ell^n(t) + 1)$. Denote by S_n this state space (restricted to the "standard" versions of the states), for each n. Denote by S_∞ the set of $[f(\cdot), \ell, r, \tau]$ such that: $f(x)$ is non-decreasing 1-Lipschitz function with periodic increments, $f(x + 1) - f(x) = \rho$, and with $f(0) = 0$; $\ell \in [0, 1)$, and $r \in [\ell, \ell + 1)$. Then it is easy to see that the state space $S \doteq S_\infty \cup [\cup_n S_n]$ is compact. This in particular implies that any sequence of distributions on S_n with $n \to \infty$ is tight, and therefore has subsequential weak limits.

Lemma 11. *Consider the model with A-perturbations. Assume $h < \rho < 1/2$. For each n consider the scaled process, with (random) initial state formed by a perturbation of an arbitrary ideal state. Fix $B \ge 1 - \rho$. Consider the stopping time T^n which is the minimum of B and the first time the process hits an ideal state. Then any subsequential (in $n \to \infty$) weak limit of the state distributions at time T^n is such that, with probability at least δ (defined in Lemma 10) the state is a MES.*

Theorem 12. *Consider the model with A-perturbations. Assume $h < \rho < 1/2$. Then, as $n \to \infty$, any subsequential weak limit of the sequence of the stationary distributions of $[f^n(\cdot, t), \ell^n(t), r^n(t), \tau^n(t)]$ is a distribution concentrated on the main equilibrium states. Consequently, the limiting flux (typical flux) $\phi(\rho) = \phi^* = (1 - \tau^*)h = p(1 - \rho)$.*

Remark 13. Theorem 12 actually holds for $h < \rho < 1$. The proof for $1/2 < \rho \le 1$ is a simplified version of that for $h < \rho < 1/2$, because when $\rho > 1/2$ a cluster always exists and A-perturbations never occur. If $\rho = 1/2$, the corresponding sequence of pre-limit systems is such that an ideal state either exists infinitely often, or does not exist infinitely often, or both. In any case, if we consider the corresponding subsequences of systems, the proof is the same as for either $h < \rho < 1/2$ or $\rho > 1/2$.

Proof of Theorem 12. For each n, consider the process with any fixed initial state at 0. Consider stopping time θ_1^n, which is the minimum of 2 and the first time (after 0) the system is in an ideal state. Consider the random system state at θ_1^n. Consider its any subsequential limit in distribution. By Lemma 7, this distributional limit is such that, almost surely, the state is either a main equilibrium state (MES) or it is an ideal state. (A small subtlety here: we use the fact that, as $n \to \infty$, the probability that a perturbation occurs immediately at a time of hitting an ideal state, vanishes.)

Consider stopping time θ_2^n, which is the minimum of $\theta_1^n + 2$ and the first time after θ_1^n the system is in an ideal state. Note that, as $n \to \infty$, the time to wait for a perturbation converges to exponentially distributed one with mean $1/\lambda$; in particular, uniformly in all large n, the probability that a time to wait for a perturbation is less than 1, is at least $\epsilon = (1 - e^{-\lambda})/2$. If the state at θ_1^n was ideal, then with probability $\geq \epsilon$, there will be a perturbation in $[\theta_1^n, \theta_1^n + 1]$. Using Lemma 11, we conclude that any subsequential limit in distribution of the system state at θ_2^n is either a MES *with probability at least* $\zeta = \epsilon\delta > 0$, or it is an ideal state. (The δ is as defined in Lemmas 10 and 11.)

We define stopping times θ_k^n for all $k \geq 2$ analogously to the definition of θ_2^n. By induction, we conclude that any distributional limit of the system state at θ_k^n is such that it is either a MES *with probability at least* $1 - (1 - \zeta)^{k-1}$, or it is an ideal state. Note that $\theta_k^n \leq 2k$. Finally, using Lemma 5(i), we establish that any subsequential limit in distribution of the system state at time $2k$ is such that with probability at least $1 - (1 - \zeta)^{k-1}$, the state is a MES. This is true for *any* fixed initial states for each n, and any k. This implies that the limit of stationary distributions of the process is concentrated on MESes. $\square$

4. I-perturbations model

Unless specified otherwise, in this section we consider the model with I-perturbations.

4.1. Some preliminary facts

As a corollary of Lemma 1, we obtain the following fact.

Lemma 14. *Suppose* $0 < \rho < h$. *Assume there are no perturbations in the (scaled) time interval* $[0, 1]$. *Then, uniformly on the initial states at time* 0, *which may contain one or multiple clusters,*

$$\lim_{n \to \infty} \mathbb{P}\{At \text{ (scaled) time } 1 \text{ the system is in an ideal state}\} = 1.$$

Proof of Lemma 14. For each n, consider an arbitrary system state at time 0, and pick any cluster (if any). Let us show that, uniformly on the initial states and the choice of the cluster,

$$\lim_{n \to \infty} \mathbb{P}\{\text{The cluster will disappear before (scaled) time } 1\} = 1. \tag{27}$$

WLOG, let the right edge of the cluster be located at (scaled) coordinate 0. Then, the "best case", in terms of the cluster *not* disappearing for as long as possible, is when this cluster contains all particles in the system. Indeed, consider the sequence of particles by going left from the right edge of the cluster. The particle may join the cluster only in this order — they cannot jump over each other. Moreover, the location at which each initially present particle joins the cluster (if the cluster does not disappear by then) is non-random and determined in advance. If all particles are initially in the cluster, this obviously minimizes their time to join the cluster (because they are already in it). Then we can use *coupling* to show that, w.p.1, the initial state with all particles in the cluster minimizes the time for the right-edge particle of the cluster to rejoin it on the left, after breaking away on the right and traveling clock-wise. (This is because after breaking away it will move unobstructed at speed 1 to the pre-determined location where it would rejoin the cluster.) We continue to observe that, w.p.1, the time for the second particle breaking away from the right to rejoin the cluster is minimized. An so on. This implies that, given the coupling, w.p.1, the time before the cluster dissolves is maximized.

From the law of large numbers, we easily verify that, for a fixed $\epsilon > 0$ such that $\rho + \epsilon < 1$, as $n \to \infty$,

$$\lim_{n \to \infty} \mathbb{P}\{\text{The cluster "dissolves" before the particle initially at 0 reaches point } -\rho - \epsilon\}$$
$$= 1.$$

This implies (27).

We are not done, because the total number of initial clusters may increase to infinity with n. However, the following observation resolves this case. Fix a small $\epsilon > 0$, such that $\rho + \epsilon < 1$. The argument we used above for a single cluster easily generalizes to show the following property:

$$\lim_{n \to \infty} \mathbb{P}\{\text{All clusters, which at time 0 overlap with } [-\epsilon, 0],$$

$$\text{will disappear before (scaled) time 1}\} = 1. \tag{28}$$

Indeed, if at time 0 we "fill in" all holes within the scaled interval $[-\epsilon, 0]$ with additional particles, we produce a single cluster (completely covering $[-\epsilon, 0]$), then the original and modified processes can be coupled so that the disappearance of the cluster in the modified system implies the disappearance of *all* clusters initially overlapping with $[-\epsilon, 0]$ in the original system. Note that the total (scaled) number of particles in the modified system is at most $\rho + \epsilon$. Therefore, (27) applies to the modified system, which proves (28). Obviously, (28) implies the lemma statement, because the circle can be covered by a finite number of ϵ-long intervals. □

From Lemma 14 we obtain the following corollary.

Lemma 15. *Consider the system with I-perturbations. Suppose $\rho < h$. Then, uniformly on all sufficiently large n and all initial process states, the probability that the process reaches an ideal state at some time $t_1 < 1$ is at least $e^{-\lambda}/2$.*

Proof. The probability that there are no perturbations in $[0, 1]$ converges to $e^{-\lambda}$. It remains to apply Lemma 14. □

Lemma 16. *Suppose $\rho < h$. Then, uniformly in n, the steady-state number of clusters is stochastically upper bounded by a proper (finite w.p.1) random variable.*

Proof. It suffices to show that the lemma statement holds for all sufficiently large n. Fix an arbitrary integer k. By Lemma 15, uniformly on all initial states at time 0, the probability that the process reaches an ideal state in $[0, k]$ is at least $1 - (1 - e^{-\lambda}/2)^k$. Considering the time interval from hitting an ideal state within $[0, k]$ until k, and taking into account the (Bernoulli) structure of the perturbation process, we easily obtain the following. (Recall also that a perturbation can increase the number of clusters by at most 2.) Uniformly on the initial states and in all large n, with probability at least $1 - (1 - e^{-\lambda}/2)^k$, the number of clusters in the system is stochastically dominated by a random variable $2(1 + H_{2\lambda k})$, where $H_{2\lambda k}$ has Poisson distribution with mean $2\lambda k$. Since this is true for all integers k, we obtain a uniform (in n) stochastic upper bound on the number of clusters in steady-state. □

4.2. HSPs and their properties

In this subsection we define HSPs for I-perturbation process, describe their properties, and give corresponding hydrodynamic limit results. For our purposes, it will suffice to consider

HSPs with a *finite number of initial clusters*. The definitions and results are described informally, because, in essence, they are straightforward generalizations of those for the A-perturbation model. We believe this informal description is sufficient — an interested reader can easily fill in all formalities.

HSP on a (finite or infinite) interval $[0, T]$ has the following structure:

$$\{f(\cdot, t), \ t \in [0, T]; \quad [\ell_i(t), r_i(t), \tau_i(t)], t \in [0, T], \ [\alpha_i, \beta_i], \quad i = 0, 1, \ldots\}. \tag{29}$$

Function $f(\cdot, t)$ describes the HSP state at t. The times $\alpha_i \leq \beta_i \leq \infty$ are the beginning and end times of the existence of i-cluster, and $\ell_i(t), r_i(t), \tau_i(t) = r_i(t) - \ell_i(t)$ are the left-end, right-end and the length of the ith cluster at time t. There is at most a finite number of initial clusters (at time 0) – they are indexed by $i = 0, 1, \ldots$ in an arbitrary order; for those initial clusters, $\alpha_i = 0$. HSP has at most a finite number of clusters by any finite time t (i.e., those with $\alpha_i \leq t$). At any given time t, the clusters present at t do not overlap. New clusters are indexed in the order of their appearance, i.e. in the order of increasing α_i. We adopt the following convention: if in a pre-limit process a cluster divides into two due to a perturbation, we "ignore" this fact and still treat the divided cluster as one; this does not cause problems, because the "holes" created within a cluster remain "invisible" in the corresponding HSP *on any finite time interval*, in the sense that their "size" remains 0 and their existence does not affect the dynamics (derivatives) of the left and right ends, $\ell_i(t)$ and $r_i(t)$, of the original cluster in the HSP. Given this convention, except possibly time 0, no more than one new cluster may appear (due to a perturbation) at any time $t > 0$. By another convention, before cluster i exists, i.e. for $t < \alpha_i$, $\ell_i(t) = r_i(t) = \tau_i(t) = 0$; and after it disappears, i.e. for $t \geq \beta_i$, $\tau_i(t) = 0$ and $\ell_i(t) = r_i(t)$ remain frozen at their value at β_i. Convergence in the HSP definition is understood as convergence to $[\alpha_i, \beta_i]$ and Skorohod convergence to $[\ell_i(t), r_i(t), \tau_i(t)], t \in [0, T]$, for each i, and the Skorohod convergence to $f(\cdot, t), \ t \in [0, T]$,.

For each cluster i, we have the (same as in the case of A-perturbations) Lipschitz properties of $\ell_i(t), r_i(t), \tau_i(t)$ within $[\alpha_i, \beta_i]$. If we denote $\tau(t) = \sum_i \tau_i$, it is still Lipschitz. We have exactly same properties describing the dynamics of each $\tau_i(t)$ in terms of the density at its left edge $\ell_i(t)$.

Finally, the following properties hold for $\rho < h$, and are proved analogously to the way it is done for A-perturbations:

(i) Denote by U_i the set of those time points in $[0, \infty)$, where $\tau_i'(t) > 0$. Then,

$$\sum_i \mathcal{L}(U_i) \leq 1/(1 + p).$$

(ii) For each cluster i, we have that $\alpha_i < \infty$ implies $\beta_i \wedge T - \alpha_i < 1$, i.e., the duration of any cluster is less than 1.

(iii) We have (18):

$$\int_0^T \tau(t)dt \leq \bar{C} = 1 + \frac{1}{2(1 + p)}.$$

4.3. I-perturbations, low-density result

Theorem 17. *Consider the model with I-perturbations and $\rho < h$. Then, as $n \to \infty$, the sequence of the stationary distributions of $[f^n(\cdot, t), \tau^n(t)]$ is such that its any subsequential limit is concentrated on the states with $\tau = 0$ and finite number of clusters. In other words, $\tau^n(\infty) \Rightarrow 0$ [equivalently, $\mathbb{E}\tau^n(\infty) \to 0$] and, consequently, the limiting flux (typical flux) $\phi(\rho) = \rho$.*

Proof of Theorem 17 is essentially same as that of Theorem 9. We do use the fact that both the pre-limit process and the HSPs have a finite number of clusters on any finite interval — this means that, just like for A-perturbations, convergence of pre-limit trajectories to an HSP implies $\tau^n(\cdot) \to \tau(\cdot)$ u.o.c. □

4.4. I-perturbations, high density result and conjecture

It is natural to conjecture that for I-perturbations, $h < \rho < 1/2$, the same result as Theorem 12 for A-perturbations holds.

Conjecture 18. *Consider the model with I-perturbations and $h < \rho < 1/2$. Then,*

$$\phi(\rho) = \phi^*.$$

The proof of Theorem 12, however does not completely go through. The key difficulty is that, for I-perturbations with $h < \rho < 1/2$, the number of clusters in steady-state is *not* uniformly stochastically bounded. The key ideas of the proof of Theorem 12 can be applied to obtain the following partial result, which shows that, when $\rho > h$, the limiting flux is strictly less than ρ.

Theorem 19. *For any $\epsilon > 0$, there exists $\delta > 0$, such that for all large n and the density $\rho \in (h + \epsilon, 1/2)$, the flux $\phi(\rho, n) < \rho - \delta$.*

We do not give a proof of this partial result, because it is a much more involved version of the proof of Theorem 12, and does not provide new insights.

Conjecture 18 is supported by extensive simulations. The simulations confirm that when $\rho > h$ the limiting flux is indeed ϕ^*. Furthermore, they show that the system "typical" state is "close to" MES in that it contains essentially a single cluster, "punctured" by holes due to perturbations. Fig. 4 contains snapshots of the states of two systems (with $n = 400$ sites in the left plot and $n = 900$ sites in the right plot) with the values $\lambda = 1$, $p = 1/2$ and $\rho = 0.4$ (so indeed $1/2 > \rho > h = p/(1 + p)$). The systems are started from a random state, where the initially occupied sites are a random sample of ρn sites out of n without replacement, and the snapshots are taken after 10^7 time slots. For each site i the plots show $x(i) * x(i + 1)$, where $x(i)$ is 1 or 0 if site i is a particle or a hole, respectively. This is a convenient way to observe clusters in a large system. One can see that there is a single cluster on the left and essentially one cluster on the right, except it is "punctured" by two holes.

5. A related zero-range model

Consider the following version (generalization) of TASEP-H. As before, there are ρn particles moving clockwise on n sites forming a circle. There are no perturbations. At each time, the following occurs. A particle that has holes as both neighbors, moves right with probability $\pi \in (0, 1]$. A particle that has another particle as the left-neighbor and a hole as the right-neighbor, moves right with probability $p \in (0, 1)$. This model, as we will see shortly, can be viewed as a *zero-range model*, considered in [7]. Denote by $v_\pi(\rho; n)$ and $\phi_\pi(\rho; n)$ the steady-state particle velocity and flux, respectively, of this system. (Subscript π will indicate that we refer to a flux for this model, as opposed to TASEP-H.) It is shown in [7] that the stationary
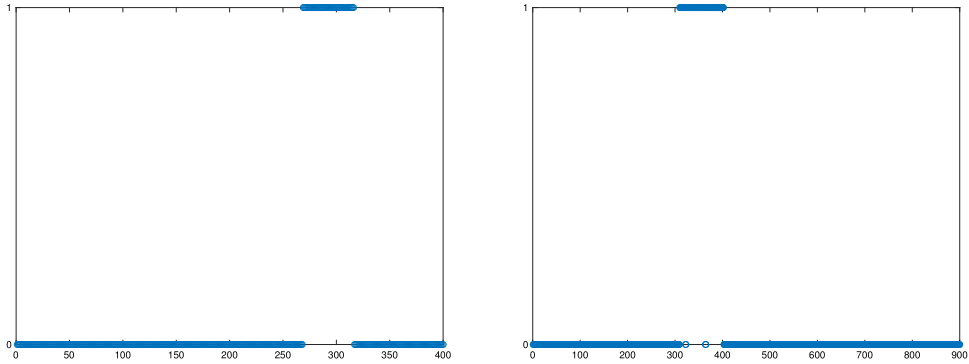
Fig. 4. The states of the system with $n = 400$ (left) and $n = 900$ (right) sites, after 10^7 time slots.

distribution of this process has a product-form; moreover, the expressions are derived, which allow to determine the limiting steady-state particle velocity and limiting flux,

$$v_\pi = v_\pi(\rho) = \lim_{n\to\infty} v_\pi(\rho; n), \quad \phi_\pi = \phi_\pi(\rho) = \lim_{n\to\infty} \phi_\pi(\rho; n),$$

in the same asymptotic regime as in this paper, with $n \to \infty$ and particle density ρ staying constant. If $\pi = 1$, we obtain exactly the basic TASEP-H model (without perturbations), so that $\phi_1 = \rho$ for any $\rho < 1/2$. If $\pi = 1 - \epsilon < 1$, but close to 1, then the zero-range model can be viewed as TASEP-H *with "perturbations" of a different kind*; namely, at each time, with small probability $\epsilon > 0$, a free particle stays in place instead of moving right. Note, however, that in the zero-range model (with π independent of n) the "perturbation" rate is much higher – $O(1)$ per particle per unit time – than for our TASEP-H with perturbations model, where the perturbation rate is $O(1/n^2)$ per particle per unit time.

Using the results of [7], we will prove the following

Theorem 20. *We have*

$$\lim_{\pi\to 1} \phi_\pi(\rho) = \begin{cases} \rho, & \rho < h, \\ \phi^* = p(1-\rho), & \rho > h. \end{cases} \tag{30}$$

In other words, $\lim_{\pi\to 1} \phi_\pi(\rho) = \phi(\rho)$, where $\phi(\rho)$ is the limiting flux of TASEP-H with A-perturbations. In particular, for $h < \rho < 1/2$, $\phi_\pi(\rho)$ is discontinuous at $\pi = 1$.

Proof. Consider the equivalent model with holes moving left: a hole with exactly one consecutive particle immediately to the left, moves left with probability π; a hole with two or more consecutive particles immediately to the left, moves left with probability p. For convenience, denote by $\gamma = 1 - \rho$ the density of holes.

Let $\eta_\pi = \eta_\pi(\rho)$ denote the limiting (in $n \to \infty$) steady-state velocity of a hole. The relation to the velocity of a particle v_π is via the equality of fluxes: $(1 - \rho)\eta_\pi = \rho v_\pi$.

It is clear that (30) it is equivalent to

$$\lim_{\pi\to 1} \phi_\pi = \begin{cases} 1 - \gamma, & \gamma > 1 - h, \\ p\gamma, & \gamma < 1 - h. \end{cases}$$

This, in turn, is equivalent to

$$\lim_{\pi \to 1} \eta_\pi = \begin{cases} \frac{1-\gamma}{\gamma}, & \gamma > 1 - h, \\ p, & \gamma < 1 - h. \end{cases}$$

Let us use notation $q = \dfrac{1 - \gamma}{\gamma}$. Note that as $h = p/(1 + p)$, the condition $\gamma > 1 - h$ is equivalent to $q < p$. So finally, we conclude that in order to show (30), it is sufficient to show that

$$\lim_{\pi \to 1} \eta_\pi = \begin{cases} q, & q < p, \\ p, & q > p, \end{cases} \tag{31}$$

or simply $\lim_{\pi \to 1} \eta_\pi = q \wedge p$.

Results of [7] are directly applicable to the model seen as movement of holes. In the notation of [7, Section 5], $p(1) = \pi$, $p(k) = p$ for $k \ge 2$. The function $f(k)$ (as in [7, Section 5]) in our case is then as follows (we can get rid of the common factor $1 - \pi$, because this function is defined up to a positive factor):

$$f(0) = 1, \quad f(1) = \frac{1}{\pi},$$
$$f(k) = \frac{1 - \pi}{\pi p} \left(\frac{1 - p}{p} \right)^{k-2}, \quad k \ge 2.$$

Note that the generating function of $\{f(k)\}$ is

$$G(z) = \sum_{k=0}^{\infty} f(k) z^k = 1 + \frac{z}{\pi} + \frac{1 - \pi}{\pi p} \frac{z^2}{1 - (\frac{1-p}{p})z}; \tag{32}$$

clearly, for real $z \in [0, p/(1 - p))$, $G(z)$ is increasing and $G(z) \uparrow \infty$, $z \uparrow p/(1 - p)$, and it is easy to check that same is true for the function $zG'(z)/G(z)$.

From [7] the (limiting in $n \to \infty$) *fugacity* $z^* = z^*(\gamma, \pi)$ satisfies the equation

$$1 - \gamma = \gamma z^* \frac{\partial \log G(z^*)}{\partial z^*} = \gamma \frac{z^* G'(z^*)}{G(z^*)}. \tag{33}$$

Note that (33) defines the fugacity z^* uniquely, and z^* is also clearly continuous w.r.t. $\pi \in (0, 1)$ and $(1 - \gamma)/\gamma \in (0, \infty)$, viewed as parameters. (We will only use the continuity in π.)

The expression for the velocity η_π in terms of fugacity z^*, is as follows (analogous to (11) in [7], holds for $z = z^*$):

$$\eta_\pi = [\pi z f(1) + p \sum_{k=2}^{\infty} z^k f(k)] G(z)^{-1}.$$

The equation above gives the dependence of η_π on z^*. For the zero-range model that we consider, this dependence is very simple, and it holds for a quite general function $p(k)$ (with $p(0) = 0$ and $0 < p(k) \le \delta < 1$ for $k \ge 1$). We have

$$f(0) = 1,$$
$$f(k) = \frac{1}{1 - p(k)} \prod_{m=1}^{k} \frac{1 - p(m)}{p(m)}, \quad k \ge 1.$$

Note that

$$p(k) f(k) = \frac{p(k)}{1 - p(k)} \prod_{m=1}^{k} \frac{1 - p(m)}{p(m)} = \prod_{m=1}^{k-1} \frac{1 - p(m)}{p(m)} = (1 - p(k-1)) f(k - 1).$$

Hence the expression for η_π is (we write z instead of z^* for simplicity below)

$$\eta_\pi = \frac{\sum_{k=1}^{\infty} p(k)f(k)z^k}{G(z)} = \frac{\sum_{k=1}^{\infty}(1 - p(k-1))f(k-1)z^k}{G(z)}$$

$$= \frac{z\sum_{k=1}^{\infty} f(k-1)z^{k-1} - z\sum_{k=1}^{\infty} f(k-1)p(k-1)z^{k-1}}{G(z)}$$

$$= \frac{zG(z) - z\eta_\pi G(z)}{G(z)} = z - z\eta_\pi.$$

Therefore,

$$\eta_\pi = \frac{z^*}{z^* + 1}, \quad \text{or} \quad z^* = \frac{\eta_\pi}{1 - \eta_\pi}. \tag{34}$$

Note that the dependence of η_π on $\pi \in (0, 1)$ (and $\gamma \in (0, 1)$) is continuous.

Using (33) and (34), for a fixed γ, we will now study the limit of η_π as $\pi \uparrow 1$, and will prove (31).

Substituting (32) into (33), and then $z^* = \eta_\pi/(1 - \eta_\pi)$, we obtain

$$(q - \eta_\pi)(p - \eta_\pi)(\eta_\pi(p - \pi - \pi p) + \pi p) = p^2(1 - \pi)\eta_\pi(1 - \eta_\pi). \tag{35}$$

Note that, as $\pi \uparrow 1$, the RHS of (35) converges to 0, *while staying positive*, because for $\pi < 1$ velocity cannot be 1, and cannot be 0 (in view of the LHS of (35)).

Note that $\eta_\pi \leq q$ must hold for any π, because q is the max velocity achievable for the holes, achieved when all particles move at speed 1. Therefore, to show that $\eta_\pi \leq q \wedge p$, it suffices to show that $\eta_\pi \leq p$ in the case $p < q$. If $\pi = p$, we have $\eta_p \leq p$ simply because the velocity of a hole in this case cannot possibly exceed p. It is also easy to see that $\eta_p < p$. (In fact, in this case the velocity is known explicitly: $\eta_p = (1 - \sqrt{1 - 4p\gamma(1-\gamma)})/(2\gamma)$, cf. [7].) As we continuously increase π in the interval $[p, 1)$, the RHS of (35) must stay positive, so for all those π we must have $\eta_\pi < p$. This completes the proof of the fact that $\eta_\pi \leq q \wedge p$ for all $\pi < 1$.

Finally, again, as we continuously increase π in the interval $[p, 1)$, velocity η_π changes continuously, the RHS of (35) must stay positive and converge to 0. The only option is that $\eta_\pi \rightarrow q \wedge p$ as $\pi \uparrow 1$. $\square$

6. Discussion and further conjectures

Our analysis of the TASEP-H model with perturbations suggests a number of generalizations and extensions. We discuss some of them.

6.1. More general perturbation rates

Consider TASEP-H with A-perturbations. Recall that for this process there are no perturbations as long as a cluster exists. As $n \rightarrow \infty$, the average time it takes the cluster in a MES to dissolve, "should" grow exponentially fast with n. (A standard large-deviations analysis should apply.) Therefore, if the inter-perturbation times grow sub-exponentially in n, then the fraction of time when the system spends in MES will dominate. This basic intuition leads to Conjecture 21.

We will say that a positive non-increasing function $g(k) \rightarrow 0$, $k \rightarrow \infty$ is sub-exponential, or write $g(k) = \text{SUBEXP}(k)$, if

$$\lim_{k \rightarrow \infty} \log(g(k))/k = 0.$$

Conjecture 21.   *Consider the model with either A-perturbations or I-perturbations. Suppose, the probability of a perturbation (under the corresponding conditions for A- and I-perturbations) at a given time is $\rho n g(n)$, where $g(n) = SUBEXP(n)$. ($g(n)$ is the perturbation rate, per particle per time unit.) Then the limiting flux (typical flux) is as follows.*
(i) If $0 < \rho < h$,

$$\phi(\rho) = \lim_{n \to \infty} \phi(\rho; n) = \rho.$$

(ii) If $h < \rho < 1/2$,

$$\phi(\rho) = \lim_{n \to \infty} \phi(\rho; n) = \phi^*.$$

The intuition that leads to Conjecture 21 for TASEP-H with perturbations, as well as the informal connection between TASEP-H and the zero-range model, also suggests that the following conjecture is very plausible.

Conjecture 22.   *Consider the zero-range model. Suppose, parameter π depends on n so that $1 - \pi(n) = SUBEXP(n)$. Then the following holds.*
(i) If $0 < \rho < h$,

$$\lim_{n \to \infty} \phi_{\pi(n)}(\rho; n) = \rho.$$

(ii) If $h < \rho < 1/2$,

$$\lim_{n \to \infty} \phi_{\pi(n)}(\rho; n) = \phi^*.$$

6.2. Slow-to-start model

Consider the following model, which is a discrete-time version of the *slow-to-start* (STS) model, studied in [4]. The model is exactly as TASEP-H, except for the following. When a cluster of particles "dissolves" – i.e., the cluster has consisted of exactly two particles, its "right" particle breaks away, and no particle joins the "left" particle from the left – the "left" particle (which now has holes on both sides) does not become free immediately; instead, at each consecutive time it moves forward and becomes free with probability p.

The STS model is primarily motivated by the movement of cars on a road. A car moves at a constant speed (normalized to 1), unless it is blocked by a stopped car ahead; when the car is unblocked, it resumes movement not immediately, but after some delay (hence, "slow start"). Paper [4] studies the continuous STS model on the real line, with particles (cars) represented by points moving continuously at constant speed, unless stopped by running into a not moving particle ahead, and with exponentially distributed restart times. The results of [4] concern the system evolution from an initial state, where none of the particles moves, towards a state where (in a sense) all particles move without bumping into each other. (Loosely speaking, the latter state is an absorbing state in the terminology of this paper.)

Considering the typical flux of the discrete-time STS model, described above, is of interest for the same reasons as for TASEP-H. The STS model with A-perturbations is defined the same way as for TASEP-H, i.e. a perturbation is applied with certain probability when all particles are free. (Recall that for STS model a particle is free when it not only has holes on both sides, but also is actually moving.) The STS model with I-perturbations is defined exactly as for TASEP-H.

It is not hard to check that the main results of the present paper for TASEP-H with A- and I-perturbations hold *as is* for the discrete-time STS model with A- and I-perturbations. Same proofs apply, with slight modifications.

Declaration of competing interest

No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work.

Appendix A. Proof of Proposition 8

The result is derived from Theorem 1.2.5 – a ballot theorem – in [11].

Without loss of generality, we can assume that $m = n$. (Otherwise, m and all k_r can be rescaled by the same factor, and this obviously does not change N.) Then, conditions (19) and (20) become, respectively,

$$\psi(j + r) - \psi(j) < r, \quad r = 1, \ldots, n, \tag{A.1}$$

and

$$N \geq \lceil n - \psi(n) \rceil. \tag{A.2}$$

For each integer $j \geq 0$, let $\eta(j) = 1$ if (A.1) holds, and $\eta(j) = 0$ otherwise. (So that, $N = \sum_0^{n-1} \eta(j)$.) Let us extend the definition of $\psi(j)$ to all real $u \geq 0$ by letting $\psi(u) = \psi(\lfloor u \rfloor)$. For each $u \geq 0$, let $\delta(u) = 1$ if

$$v - u \geq \psi(v) - \psi(u) \quad \text{for all } v \geq u, \tag{A.3}$$

and $\delta(u) = 0$ otherwise. Denote by D the set of those $u \in [0, n)$, for which $\delta(u) = 1$. By Theorem 1.2.5 in [11], $\mathcal{L}(D) = n - \psi(n)$, where $\mathcal{L}$ denotes the Lebesgue measure. For any interval $[j, j + 1)$, $j = 0, 1, \ldots, n - 1$, observe the following: if $\delta(u) = 1$ for some $j < u < j + 1$, then necessarily $\eta(j) = 1$. Indeed, in this case, for any $r = 1, \ldots, n$,

$$r - [\psi(j + r) - \psi(j)] = (u - j) + (j + r - u) - [\psi(j + r) - \psi(u)] \geq u - j > 0,$$

where the first inequality is by (A.3). This implies that for each interval $[j, j + 1)$ such that $\mathcal{L}(D \cap [j, j + 1)) > 0$ we have $\eta(j) = 1$. Since, obviously, $\mathcal{L}(D \cap [j, j + 1)) \leq 1$ for any j, and

$$\sum_{j=0}^{n-1} \mathcal{L}(D \cap [j, j + 1)) = \mathcal{L}(D) = n - \psi(n),$$

we see that the (integer) number N of those $j \in \{0, 1, \ldots, n - 1\}$ for which $\eta(j) = 1$ cannot be less than $n - \psi(n)$. This proves (20). $\quad \square$

Appendix B. A CSMA model indirectly motivating the holdback property of TASEP-H

In this section we describe an interacting particle system, which is a model of a fairly realistic wireless network under a CSMA protocol. This system is far more complicated than TASEP-H, but exhibits qualitatively similar behavior. We will formulate some hypotheses and discuss the difficulties one faces in analyzing such a system. We briefly introduce the model here and refer to [10] for an account of known results on the topic, in particular stability analysis.

As in the TASEP-H setting, there are n sites arranged in a circle, and ρn particles, moving clockwise ("right") in discrete time. (A time index is sometimes referred to as a "time slot".) The particle density $\rho \in (0, \infty)$ (not necessarily less than 1). At each time a particle may move to its right-neighbor site. There is no restriction on the number of particles present at a site at a
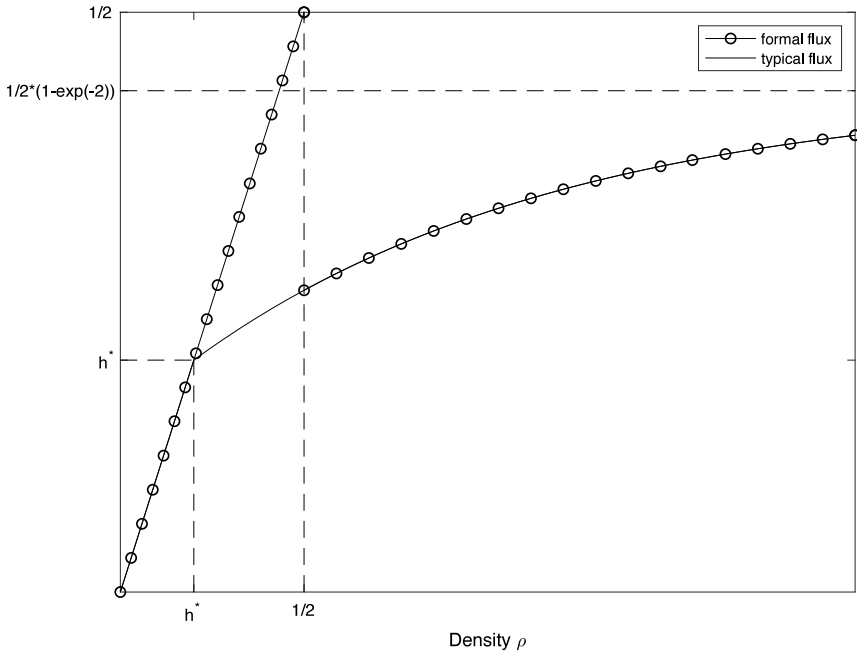
Fig. B.5. Formal and typical flux in the CSMA model, versus density ρ, as suggested by simulations. For $\rho < h^*$ and $\rho > 1/2$ the formal and typical fluxes coincide.

given time. A particle may move to the right-neighbor site even if that site has other particles. The particles that do move at a given time are chosen according to a randomized competition described as follows. At each time the sites are given access priorities, forming a permutation of numbers $1, \ldots, n$, picked independently (across time slots), uniformly at random from all possible permutations. From the site with the highest priority, if it is not empty (has one or more particles), exactly one particle (chosen at random, say) moves to its right-neighbor site. From the site with the second-highest priority, exactly one particle moves to the right-neighbor site, as long as the site is not empty and that no particle has already moved from any of its two neighbor sites. And so on until all sites are checked in their priority order. The procedure described above is repeated independently over time slots.

The restriction that a particle can only move from a site if no particle moved from any of the neighbors of the site models the interference in wireless networks, and it is the most important characteristic of models of these networks.

We are interested in the flux of this system. As in the TASEP-H model, if $\rho < 1/2$ an ideal (absorbing) state will be eventually reached and the system will stay in it, thus the formal flux is equal to ρ as long as $\rho < 1/2$ (see Fig. B.5).

A typical flux is defined analogously to that for TASEP-H, i.e. as the limit of the flux of the system with perturbations, as $n \to \infty$. The typical flux that we observe in simulations is shown in Fig. B.5. (The typical flux plot in Fig. B.5 is just a qualitative illustration — it does not describe a specific simulation experiment.) We observe that the typical flux is monotone increasing, asymptotically converging to the *parking constant* $c = (1/2)(1 - e^{-2}) \approx 0.43$ as $\rho \to \infty$. The asymptotic limit c is intuitive. Indeed, as $\rho \to \infty$, almost all sites will

be occupied most of the time. If so, when n is large, the fraction of sites that will move a particle at a given time will be equal to the expected number of cars parked per space-slot in the so-called discrete parking system (we refer to [10] for an explanation of the connections between the models). We also *conjecture* that there is a phase transition at some density level h^*. The conjectured phase transition should be qualitatively similar to that for TASEP-H model. Namely, when the density $\rho < h^*$, the clusters that may emerge will not have a tendency to grow very large, because, if/when they become large they "lose" particles on the right faster than "acquire" new particle on the left (due to free particles joining it). As a result, when $\rho < h^*$ almost all (in the $n \to \infty$ limit — all) particles remain free. If $\rho > h^*$, very large clusters form and contain a non-zero fraction of particles, while the density of particles in the sparse intervals is h^*. (The value of h^* is not known to us. A simple heuristic argument leads us to believe that $h^* \geq 1/9$. Indeed, the rate at which particles "break away" on the right from a cluster, followed by a sparse interval, is at least $1/8$; within any two time slots, a particle breaks away with probability at least $(1/2)^2$. And if the density of the sparse interval on the left of the cluster is less than $1/9$, then particles join the cluster on the left at the rate at most $(1/9)/(1 - 1/9) = 1/8$. Thus, if the overall particle density $\rho < 1/9$, clusters should have a tendency to dissolve.)

We further conjecture that for $\rho > 1/2$ the formal flux coincides with the typical flux. That is, the formal flux has a negative "jump" at $\rho = 1/2$, and then is equal to the typical flux for $\rho > 1/2$ (see Fig. B.5).

We note that an analysis of the system in this section is much more difficult than that of the TASEP-H system due to a number of reasons, including, but not limited to, the following: at any time, any cluster may in principle break into any number of clusters even without perturbations; similarly, any number of clusters may merge and form a much larger cluster; even if the rightmost site of a cluster moves a particle to the right, the latter particle does not necessarily break away from the cluster, because the site may have had other particles in it.

References

[1] R.A. Blythe, M.R. Evans, Nonequilibrium steady states of matrix-product form: a solver's guide, J. Phys. A 40 (2007) R333–R441.

[2] A. Borodin, I. Corwin, Discrete time q-taseps, Int. Math. Res. Not. 2015 (2) (2015) 499–537.

[3] A. Borodin, I. Corwin, T. Sasamoto, From duality to determinants for q-tasep and asep, Ann. Probab. 42 (6) (2014) 2314–2382.

[4] F.C. Cáceres, P.A. Ferrari, E. Pechersky, A slow-to-start traffic model related to a m/m/1 queue, J. Stat. Mech. Theory Exp. 2007 (2007) P07008.

[5] D. Chowdhury, L. Santen, A. Schadschneider, Statistical physics of vehicular traffic and some related systems, Phys. Rep. 329 (4–6) (2000) 199–329.

[6] M. Evans, Bose-einstein condensation in disordered exclusion models and relation to traffic flow, Europhys. Lett. 36 (1) (1996) 13–18.

[7] M. Evans, Exact steady states of disordered hopping particle models with parallel and ordered sequential dynamics, J. Phys. A: Math. Gen. 30 (1997) 5669–5685.

[8] D. Helbing, Traffic and related self-driven many-particle systems, Rev. Modern Phys. 73 (4) (2001) 1067–1141.

[9] T.M. Liggett, Interacting Particle Systems, Springer Science & Business Media, 2012.

[10] S. Shneer, A. Stolyar, Stability conditions for a discrete-time decentralised medium access algorithm, Ann. Appl. Probab. 28 (6) (2018) 3600–3628.

[11] L. Takacs, Combinatorial Methods in the Theory of Stochastic Processes, John Wiley and Sons, New York-London-Sydney, 1967.