LARGE-SCALE BEHAVIOR OF A PARTICLE SYSTEM WITH MEAN-FIELD INTERACTION: TRAVELING WAVE SOLUTIONS

ALEXANDER STOLYAR ⓘ,* *University of Illinois at Urbana-Champaign*

Abstract

We use probabilistic methods to study properties of mean-field models, which arise as large-scale limits of certain particle systems with mean-field interaction. The underlying particle system is such that n particles move forward on the real line. Specifically, each particle 'jumps forward' at some time points, with the instantaneous rate of jumps given by a decreasing function of the particle's location quantile within the overall distribution of particle locations. A mean-field model describes the evolution of the particles' distribution when n is large. It is essentially a solution to an integro-differential equation within a certain class. Our main results concern the existence and uniqueness of—and attraction to—mean-field models which are traveling waves, under general conditions on the jump-rate function and the jump-size distribution.

Keywords: Particle system; mean-field model dynamics; asymptotic behavior; distributed system synchronization

2020 Mathematics Subject Classification: Primary 90B15
Secondary 60K25; 60K35

1. Introduction

In this paper we use probabilistic methods to study properties of mean-field models, which describe large-scale behavior of certain particle systems with mean-field interaction. A mean-field model is essentially a solution to an integro-differential equation within a certain class. Our focus is on the existence and uniqueness of—and attraction to—mean-field models which are traveling waves.

1.1. A particle system giving rise to the mean-field model

The basic particle system which gives rise to our mean-field model is as follows. (This system—or, rather, a special case of it—was first introduced and studied in [6, 7].) There are n particles, moving in the positive direction ('right') on the real axis $\mathbb{R}$. Each particle moves in jumps, as follows. For each particle there is an independent Poisson process of rate $\mu > 0$. At the time points of this Poisson process the particle jumps to the right with probability $\eta_n(\nu)$, where ν is its quantile in the current empirical distribution of the particles' locations; that is, $\nu = \ell/n$ when the particle location is ℓth from the left. With complementary probability $1 - \eta_n(\nu)$ the particle does not jump. To have the model well-defined, assume that quantile-ties between co-located particles are broken uniformly at random. Assume that,

for each n, $\eta_n(\nu)$, $0 \le \nu \le 1$, is non-increasing, and that, as $n \to \infty$, it uniformly converges to a continuous, strictly decreasing function $\eta(\nu)$, $0 \le \nu \le 1$, with $\eta(0) = 1$, $\eta(1) = 0$. The jump sizes, when a particle does jump, are given by independent and identically distributed (i.i.d.) non-negative random variables with cumulative distribution function $J(y)$, $y \ge 0$; we denote by $\bar{J}(y) = 1 - J(y)$ the complementary cumulative distribution function. In this paper we assume that for some integer $\ell \ge 2$, the jump size distribution has finite ℓth moment, and we define

$$m^{(k)} \doteq \int_0^\infty y^k dJ(y) < \infty, \qquad k = 1, 2, \ldots, \ell. \tag{1.1}$$

So, $m^{(1)} < \infty$ is the mean jump size.

Let $f^n(x, t)$ be the (random) empirical distribution of the particle locations at time t; that is, $f^n(x, t)$ is the fraction of particles located in $(-\infty, x]$ at time t. As $n \to \infty$, it is very intuitive that $f^n(x, t)$ converges (in an appropriate sense, under appropriate conditions) to a deterministic function $f(x, t)$ such that $f(\cdot, t)$ is a distribution function for each t, and the following equation holds:

$$\frac{\partial}{\partial t} f(x, t) = -\mu \int_{-\infty}^x d_y f(y, t) \eta(f(y, t)) \bar{J}(x - y), \tag{1.2}$$

where d_y means the differential in y.

At this point let us describe only the intuition for (1.2). For each t, the distribution $f(\cdot, t)$ approximates the distribution of particles $f^n(\cdot, t)$ when n is large. Since particles move right, $f(x, t)$ is non-increasing in t for each x. So $(\partial/\partial t) f(x, t) \le 0$, and its value should be equal to the right-hand side of (1.2), which gives the instantaneous rate (scaled by $1/n$ and taken with a minus sign) at which particles jump over the point x at time t. We will call a function $f(x, t)$ satisfying (1.2) a *mean-field model*. The formal meaning of (1.2) and the definition of a mean-field model will be given later.

It is also intuitively clear (and easy to make formal, as we do later) that, for any mean-field model, the speed at which the mean $\int x d_x f(x, t)$ of the distribution $f(\cdot, t)$ moves right must be equal to $v = m^{(1)} \mu \int_0^1 \eta(\nu) d\nu$. Suppose a mean-field model $f(x, t)$ that is a traveling wave exists, namely, $f(x, t) = \phi(x - vt)$ for some distribution function $\phi(\cdot)$ which we will call a *traveling wave shape*. By substituting into (1.2), we see that any traveling wave shape ϕ must satisfy the equation

$$v \phi'(x) = \mu \int_{-\infty}^x \phi'(y) \eta(\phi(y)) \bar{J}(x - y) dy. \tag{1.3}$$

We will make this statement formal later.

1.2. Motivation for the particle system

The original motivation for the particle system described above is an idealized model [6, 7] of distributed parallel simulation. The n particles represent n components, or sites, of one large system. Each site is being simulated by a separate computer (processor). The sites are interdependent, so their simulations cannot run independently. A particle location at a given real time represents the 'local simulation time' of the corresponding site, which is the time in the simulated system up to which the simulation of this site is valid. After a site's independent simulation runs for some real time, the site tries to update its state and advance its local time. However, this local time advance is not always possible, because the site's evolution depends on a number of other sites, whose local times may be lagging behind; if the site's local time cannot be advanced, it will 'go back' and start simulation again from the current local time.

In general, a site's local time advance is more likely to occur if it is 'further behind' the local times of other sites. The model in [6, 7] specifically assumes the following: each particle gets 'urges' to jump forward (advance local time) as an independent Poisson process of rate μ; when a particle does get a jump urge, it actually jumps only if K other particles, $K \geq 1$, chosen uniformly at random, are currently ahead of it. If ν is the quantile of the particle location, then the probability of the particle jump, $\eta_n(\nu)$, is such that $\eta_n(\nu) \to \eta(\nu) = (1 - \nu)^K$ as $n \to \infty$. It is also assumed in [6, 7] that the jump size distribution $J(\cdot)$ is exponential. The paper [6] further assumes that the system's initial state is such that the locations of the n particles are drawn independently from a given distribution. We will comment on these additional assumptions in detail later, in Section 3.3.

1.3. Prior results on the particle system and its mean-field model

The papers [6, 7] address two different issues related to the particle system and solutions of (1.2):

- *Limit transition from $f^n(x, t)$ to a mean-field model $f(x, t)$, as $n \to \infty$.*
  In [6] it is proved, under certain additional assumptions, that, as $n \to \infty$, the random process $f^n(x, t)$ indeed converges to a deterministic process $f(x, t)$ satisfying (1.2).

- *Convergence of a mean-field model $f(x, t)$ to a traveling wave solution, as $t \to \infty$.* In [7] the following is proved, under certain additional assumptions. *If a traveling wave shape ϕ, i.e. a solution to (1.3), exists*, then it is unique (up to a shift) and, as $t \to \infty$, a mean-field model $f(x, t)$ converges to the traveling wave solution, that is, $f(\cdot + vt, t) \to \phi(\cdot)$. The question of the existence of a traveling wave shape ϕ is left open in [7], except for the case of an exponential distribution $J(\cdot)$ and $\eta(\nu) = (1 - \nu)^K$, when (1.3) is easily explicitly solvable.

A detailed discussion of the results of [6, 7] is given in Section 3.3, which also discusses other related work.

1.4. Main results of this paper

We study the properties of mean-field models, specifically the existence and uniqueness of, and convergence to, traveling wave shapes. The convergence to a traveling wave shape is important, because if it holds, it means that the particles' locations 'remain close to each other' (as they move right) *regardless of the number of particles,* in the sense that the distribution of the particles' locations stays close to a certain shape, which moves right at the constant speed v.

Our main results are the following:

- We prove (in Theorem 3.1) the existence of a traveling wave shape ϕ, for general jump size distribution $J(\cdot)$ and general (strictly decreasing continuous) $\eta(\cdot)$. Moreover, as a distribution, ϕ has finite $(\ell - 1)$th moment, $\int_y |y|^{\ell - 1} d\phi(y) < \infty$, for the $\ell \geq 2$ in (1.1).

- Under the additional condition that $J(\cdot)$ has positive density (bounded away from 0 on compact sets), we show (in Theorem 3.2) the uniqueness (up to a shift) of the traveling wave shape ϕ and convergence to it, $f(\cdot + vt, t) \to \phi(\cdot)$, for any mean-field model.

- As the main tool for analysis, we introduce and study the properties of traveling wave shapes within 'finite frames'. The existence of a traveling wave shape is then proved by letting the frame size go to infinity. These results may be of independent interest.

We emphasize that the results of this paper concern properties of formally defined mean-field models. The question of whether/when the convergence to a mean-field model holds is a separate issue, partially addressed in [6], as pointed out above in Section 1.3; it will be discussed in more detail in Section 3.3.1.

1.5. Overview of the technical approach and key challenges

Our approach to proving the existence of a traveling wave shape (Theorem 3.1) is as follows. A traveling wave shape ϕ can be characterized as a fixed point of an operator A, which maps a probability distribution ϕ (describing a distribution of particle locations) into the stationary distribution $A\phi$ of a single particle, evolving within the 'environment' given by ϕ. (This is done in Section 4.) Specifically, the particle jumps right as described in our motivating particle system, except its current location quantile is that within the distribution ϕ; between jumps the particle moves left at the constant speed v. Since ϕ is a fixed point, it is natural to try to obtain its existence using in some way the Brouwer fixed point theorem. In our case this cannot be done directly, because the space of proper distributions on the real line in non-compact.

To address this challenge, we introduce and study traveling wave shapes within 'finite frames'—they are fixed points of finite-frame versions of the operator A, and play a key role in our analysis. (This is done in Section 5.) A finite-frame version $A^{(w;B)}$ of the operator A maps a distribution γ, concentrated on the compact segment (finite frame) $[-B, B]$, into the stationary distribution $A^{(w;B)}\gamma$ of a single particle, which evolves within the environment given by γ, and whose movement is restricted to $[-B, B]$ as follows: the particle cannot 'jump over B'; between jumps it moves left at the constant speed w, but cannot move to the left of $-B$. We show that operator $A^{(w;B)}\gamma$ is continuous in γ, as well as in the parameters B and w. Given the compactness of a finite frame, the existence of fixed points of $A^{(w;B)}$—finite-frame traveling wave shapes—follows. We further show that a finite-frame traveling wave shape is unique and depends continuously on the parameters B and w; showing the uniqueness (Lemma 5.3(i)) is the most involved part here.

We then let $B \to \infty$ along some sequence, and for each B we choose $w = w(B)$ such that the corresponding finite-frame traveling wave shape γ_B has its median exactly at 0. The rest of the proof of the existence of a traveling wave shape (in Section 6.1) shows that the family of distributions $\{\gamma_B\}$ is tight, and any subsequential limit ϕ must be a traveling wave shape. Proving the tightness is the key technical part of the paper. It involves, first, showing (Lemma 6.1) that, necessarily, $w \to v$ as $B \to \infty$; if not, the particle process would have inherent positive or negative drift, and the proof obtains a contradiction with the fact that the median of γ_B stays at 0. Finally, Lemma 6.2 shows the tightness of $\{\gamma_B\}$ itself; if not, the proof considers the single-particle process under a space as well as a space/time rescaling and shows that, for large B, these rescaled processes would have a strictly negative steady-state drift of a quadratic Lyapunov function—a contradiction. The use of a quadratic Lyapunov function here requires that the jump size distribution has finite second moment; this is where the finite-second-moment assumption is used in the existence proof.

1.6. Outline of the rest of the paper

Section 2 gives some basic notation used throughout the paper. In Section 3 we formally define mean-field models, state our main results, and discuss previous related work. The results in Section 4 characterize a traveling wave shape as a fixed point of an operator A; here we also obtain the finiteness of moments of $A\phi$. In Section 5 we introduce and study traveling wave

shapes within finite frames. Section 6 contains the proofs of our main results, Theorems 3.1 and 3.2. The discussion in Section 7 includes a generalization of our main results; possible relaxation of assumptions; and a conjecture about the limit of the stationary distributions of the particle system, as the number of particles $n \to \infty$.

2. Basic notation

The sets of real and non-negative real numbers are denoted by $\mathbb{R}$ and $\mathbb{R}_+$, respectively. As a measurable space, $\mathbb{R}$ is endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$. Convergence to a set, $x \to S \subseteq \mathbb{R}$, means $\inf_{y \in S} |x - y| \to 0$. For $x \in \mathbb{R}$, $\lfloor x \rfloor$ is the largest integer not greater than x. For a condition/event A, $I\{A\} = 1$ if A holds, and $I\{A\} = 0$ otherwise.

For functions $h(x)$ of a real x, $h(x+)$ and $h(x-)$ are the right and left limits; $(d/dx)h(x) = h'(x)$ is the derivative; $(d^+/dx)h(x)$ and $(d^-/dx)h(x)$ are the right and left derivatives;

$$\frac{d_\ell^+}{dx}h(x) = \liminf_{y \downarrow x} \frac{h(y) - h(x)}{y - x}, \qquad \frac{d_u^+}{dx}h(x) = \limsup_{y \downarrow x} \frac{h(y) - h(x)}{y - x}$$

are the lower and upper right derivative numbers; $\|h\| = \sup_x |h(x)|$ is the sup-norm, and $\overset{u}{\to}$ is the corresponding uniform convergence; $\|h\|_1 = \int_x |h(x)| dx$ is the L_1-norm, and $\overset{L_1}{\to}$ is the corresponding convergence; $h \overset{u.o.c.}{\to} g$ means convergence that is *uniform on compact sets*; we denote by θ_c, $c \in \mathbb{R}$, the shift operator $\theta_c h(x) = h(x - c)$; $h(x)$ is called c-Lipschitz if it is Lipschitz with constant $c \geq 0$. The notation $d_x h(x, t)$ for a multivariate function means the differential in the variable x.

We say that a function $g(x)$ of a real x is RCLL if it is right-continuous with left limits. The domain of $g(x)$ will be clear from the context; usually, $x \in \mathbb{R}$. For RCLL functions, $\overset{J_1}{\to}$ denotes Skorokhod (J_1) convergence (cf. [5]).

For non-decreasing RCLL functions, $h \overset{w}{\to} \gamma$ denotes weak convergence, namely, convergence at every point of continuity of γ. The symbol $\overset{w}{\to}$ is also used more generally, to denote weak convergence of measures.

A non-decreasing RCLL function $\gamma = (\gamma(x), x \in \mathbb{R})$ is a probability distribution function if $\lim_{x \downarrow -\infty} \gamma(x) \geq 0$ and $\lim_{x \uparrow \infty} \gamma(x) \leq 1$; a probability distribution function γ is *proper* if $\lim_{x \downarrow -\infty} \gamma(x) = 0$ and $\lim_{x \uparrow \infty} \gamma(x) = 1$; thus, an improper γ may have atoms at $-\infty$ and/or ∞. We use the terms *probability distribution function*, *distribution function*, and *distribution* interchangeably. Unless explicitly stated otherwise, a distribution means a proper distribution. The inverse (νth quantile) of a (proper or improper) distribution γ is $\gamma^{-1}(\nu) \doteq \inf\{y \mid \gamma(y) \geq \nu\}$, $\nu \in [0, 1]$; $\gamma^{-1}(\nu) = \infty$ when the set under inf is empty. We use a usual stochastic order (dominance) relation between probability distributions on $\mathbb{R}$: $g \preceq \gamma$ if and only if $g(x) \geq \gamma(x)$, $x \in \mathbb{R}$; we refer to this as γ dominating g. For a distribution γ and a function h, we write $\gamma h \doteq \int_{\mathbb{R}} h(y) d\gamma(y)$,

$$\int_{x-}^\infty h(y) d\gamma(y) \doteq \int_{[x, \infty]} h(y) d\gamma(y), \qquad \int_{x+}^\infty h(y) d\gamma(y) \doteq \int_{(x, \infty]} h(y) d\gamma(y).$$

When G_k, G are operators mapping a function of $x \in \mathbb{R}$ into another function of $x \in \mathbb{R}$, the convergence $G_k h \to Gh$, or $\lim G_k h = Gh$, always means uniform convergence $\|G_k h - Gh\| \to 0$.

Suppose we have a Markov process taking values in $\mathbb{R}$, with $P^t(x, H)$, $t \geq 0$, $x \in \mathbb{R}$, $H \in \mathcal{B}(\mathbb{R})$, being its transition function. As an operator, P^t is given by $P^t h(x) \doteq \int_y P^t(x, dy)h(y)$; $I = P^0$ is the identity operator. The (infinitesimal) generator G of the process is

$$Gh \doteq \lim_{t \downarrow 0}(1/t)[P^t - I]h.$$

A function h is within the domain of the generator G if Gh is well-defined.

We will also use the following non-standard notation throughout the paper. For a probability distribution function γ, a strictly decreasing continuous function $\eta(\nu)$, $0 \leq \nu \leq 1$, and $y \in \mathbb{R}$, we define

$$\bar{\eta}(y, \gamma) \doteq \begin{cases} \eta(\gamma(y)) & \text{when } \gamma(\cdot) \text{ is continuous at the point } y, \\ (\nu_2 - \nu_1)^{-1} \int_{\nu_1}^{\nu_2} \eta(\nu)d\nu & \text{otherwise,} \end{cases}$$

where $\nu_2 = \gamma(y)$, $\nu_1 = \gamma(y-)$.

Convergence in probability is denoted by $\xrightarrow{P}$, and 'w.p.1' means *with probability* 1.

3. Mean-field model of a large-scale system

3.1. Mean-field model definition

Let $\eta(\nu)$, $0 \leq \nu \leq 1$, be a continuous, strictly decreasing function, with $\eta(0) = 1$, $\eta(1) = 0$. Let $J(y)$, $y \geq 0$, be the cumulative distribution function of a probability distribution, concentrated on $\mathbb{R}_+$ and satisfying (1.1) for some integer $\ell \geq 2$; let $\bar{J}(y) = 1 - J(y)$. We now introduce the following definition.

Definition 3.1. A function $f(x, t)$, $x \in \mathbb{R}$, $t \in \mathbb{R}_+$, will be called a *mean-field model* if it satisfies the following conditions:

(a) For any t, as a function of x, $f(x, t)$ is a probability distribution function; that is $f(\cdot, t)$ is non-decreasing and RCLL, with $\lim_{x \to -\infty} f(x, t) = 0$ and $\lim_{x \to \infty} f(x, t) = 1$.

(b) For any x, $f(x, t)$ is non-increasing and c-Lipschitz in t, with constant c independent of x.

(c) For any x, for any t where the partial derivative $(\partial/\partial t)f(x, t)$ exists (which is almost all t with respect to Lebesgue measure, by the Lipschitz property), the equation

$$\frac{\partial}{\partial t} f(x, t) = -\mu \int_{-\infty}^{x} d_y f(y, t)\bar{\eta}(y, f(\cdot, t))\bar{J}(x - y), \tag{3.1}$$

holds.

Equation (3.1) is a more general form of (1.2), allowing $f(x, t)$ to be RCLL in x, rather than continuous. If $f(\cdot, t)$ is continuous at y, then $\bar{\eta}(y, f(\cdot, t)) = \eta(f(y, t))$; if $f(\cdot, t)$ has a jump at y then $\bar{\eta}(y, f(\cdot, t))$ is $\eta(\nu)$ averaged over $\nu \in [f(y-, t), f(y, t)]$.

Note that the following holds for any mean-field model $f(x, t)$. Let

$$v \doteq \mu m^{(1)} \int_0^1 \eta(\nu)d\nu.$$

Then for any $\tau \leq t$,

$$\int_{-\infty}^{\infty} [f(x, \tau) - f(x, t)]dx = v(t - \tau). \tag{3.2}$$

Indeed, if we denote by $h(x, t)$ the right-hand side of (3.1), we obtain

$$
\begin{aligned}
\|h(\cdot, t)\|_1 &= \int_{-\infty}^{\infty} dx\, \mu \int_{-\infty}^{x} d_y f(y, t) \bar{\eta}(y, f(\cdot, t)) \bar{J}(x - y) \\
&= \int_{-\infty}^{\infty} dx\, \mu \int_{0}^{f(x,t)} dv\, \eta(v) \bar{J}(x - f^{-1}(v, t)) \\
&= \mu \int_{0}^{1} \eta(v) dv \int_{f^{-1}(v,t)}^{\infty} dx\, \bar{J}(x - f^{-1}(v, t)) \\
&= \mu \int_{0}^{1} \eta(v) dv \int_{0}^{\infty} \bar{J}(\xi) d\xi = \mu m^{(1)} \int_{0}^{1} \eta(v) dv = v,
\end{aligned}
$$

where $f^{-1}(v, t)$ is the inverse of $f(y, t)$ with respect to y. For any x, since $f(x, t)$ is Lipschitz—and then absolutely continuous—in t, we have

$$
f(x, \tau) - f(x, t) = \int_{\tau}^{t} [-h(x, \xi)] d\xi.
$$

Then

$$
\begin{aligned}
\int_{-\infty}^{\infty} [f(x, \tau) - f(x, t)] dx &= \int_{-\infty}^{\infty} \int_{\tau}^{t} [-h(x, \xi)] d\xi\, dx \\
&= \int_{\tau}^{t} \int_{-\infty}^{\infty} [-h(x, \xi)] dx\, d\xi = \int_{\tau}^{t} \|h(\cdot, \xi)\|_1 d\xi = v(t - \tau).
\end{aligned}
$$

The equality ('conservation law') (3.2) implies in particular that, if the mean of the distribution $f(\cdot, \tau)$,

$$
\bar{f}(\tau) \doteq \int_{-\infty}^{\infty} x d_x f(x, \tau),
$$

is well-defined and finite (i.e., $\int_{-\infty}^{\infty} |x| d_x f(x, \tau) < \infty$), then $\bar{f}(t)$ is finite for all $t \geq \tau$, and

$$
\bar{f}(t) - \bar{f}(\tau) = v(t - \tau).
$$

In other words, if $\bar{f}(\tau)$ is finite, then $\bar{f}(t)$ is finite for all $t \geq \tau$, and it moves right at the constant speed v.

Definition 3.2. Suppose a mean-field model *f(x, t)* which is a traveling wave exists; namely, $f(x, t) = \phi(x - vt)$ for some probability distribution function $\phi(\cdot)$. Such a function ϕ will be called a *traveling wave shape*.

Note that by (3.2) the speed of a traveling wave can only be v. It is straightforward to see that *a function ϕ is a traveling wave shape if and only if it is a Lipschitz continuous distribution function, satisfying* (1.3) *at any x where $\phi'(x)$ exists (which is almost any x).* Furthermore, from the form of (1.3) we see that, in fact, *ϕ satisfies* (1.3) *for each x, and the derivative $\phi'(x)$ is continuous.*

Note that if $\phi(x)$ is a traveling wave shape, then so is $\phi(x - c)$ for any constant real shift c.

3.2. Main results

The following is the main result of this paper. It proves the existence of a traveling wave shape ϕ, under the general assumptions of our model. It also shows that ϕ has a finite $(\ell - 1)$th moment.

Theorem 3.1. *Assume* (1.1). *Then the following hold:*

(i) *There exists a traveling wave shape $\phi(\cdot)$.*

(ii) *Any traveling wave shape $\phi(\cdot)$ is such that, for the integer $\ell \geq 2$ in the condition* (1.1),

$$\int_0^\infty |y|^{\ell-1} d\phi(y) < \infty. \tag{3.3}$$

We also obtain the following uniqueness and convergence result. It is proved by combining Theorem 3.1 with the results of [7].

Theorem 3.2. *Assume, in addition, that the jump size distribution is such that*

$$J(\cdot) \text{ has density } J'(y) > 0, \text{ bounded away from } 0 \text{ on compact subsets of } \mathbb{R}_+. \tag{3.4}$$

Then the following hold:

(i) *The traveling wave shape $\phi(\cdot)$ is unique, up to a shift $\phi(\cdot - c)$.*

(ii) *If a mean-field model $f(x, t)$ is such that the initial mean $\bar{f}(0)$ is well-defined and finite, then the convergence to the unique traveling wave shape ϕ takes place,*

$$f(\cdot + vt, t) \xrightarrow{L_1} \phi(\cdot) \quad \text{and} \quad f(\cdot + vt, t) \xrightarrow{u} \phi(\cdot), \tag{3.5}$$

where ϕ is uniquely centered by the condition $\int_y y d\phi(y) = \bar{f}(0)$.

Note that to have a 'clean' convergence to the traveling wave shape ϕ, as in (3.5), *some* additional conditions on the distribution $J(\cdot)$ are required. For example, if $f(\cdot, 0)$ is concentrated on a lattice $\{ck, k \text{ is integer}\}$, and the distribution $J(\cdot)$ is arithmetic, concentrated on $\{ck, k = 1, 2, \ldots\}$, then the convergence (3.5) is impossible, even though a traveling wave shape ϕ does exist (by Theorem 3.1) and might be unique.

Most of the rest of the paper is devoted to the proof of Theorem 3.2. The proof constructs a traveling wave shape ϕ as a limit of traveling wave shapes within finite frames. The finite-frame traveling wave shapes and their analysis may be of independent interest.

Before proceeding with the proofs, we observe that, without loss of generality, we can assume that

$$\mu = 1, \qquad m^{(1)} = 1. \tag{3.6}$$

Indeed, any mean-field model can be reduced to the corresponding model satisfying (3.6), by time and space rescaling. So, in all the proofs we do assume (3.6), in which case

$$v = \int_0^1 \eta(\xi) d\xi,$$

(3.1) becomes

$$\frac{\partial}{\partial t} f(x, t) = -\int_{-\infty}^x d_y f(y, t) \eta(f(y, t)) \bar{J}(x - y), \tag{3.7}$$

and (1.3) becomes

$$v\phi'(x) = \int_{-\infty}^{x} \phi'(y)\eta(\phi(y))\bar{J}(x-y)dy. \tag{3.8}$$

3.3. Previous work

3.3.1. *Detailed discussion of* [6]. As mentioned earlier, the paper [6] proves that, as $n \to \infty$, the random process $f^n(x, t)$ converges to a mean-field model $f(x, t)$, *under the additional assumptions* that we now discuss.

The first additional assumption is on *the form of the jump probability functions $\eta_n(\cdot)$ in the pre-limit particle system.* The paper considers the specific particle system in which, when a particle gets a jump urge, it actually jumps only if K other particles, $K \geq 1$, chosen uniformly at random, are ahead of it. In this case $\eta_n(v) \to \eta(v) = (1-v)^K$ (see Section 1.2). The important part of this assumption about $\eta_n(v)$ is that the decision as to whether or not the particle jumps depends only on the locations of a set of other particles chosen uniformly at random, and the cardinality of that random set is bounded above by some K. Consequently, the jump probability assumption in [6] can be generalized, for example, as follows: a particle that gets an urge to jump chooses K other particles uniformly at random, and its decision to jump or not depends only on k/K, where k is the number of chosen particles that are behind this particle. Under this more general assumption on $\eta_n(v)$, the results of [6] still apply, as long as $\eta_n(\cdot)$ uniformly converges to some strictly decreasing continuous function $\eta(\cdot)$, with $\eta(0) = 1$ and $\eta(1) = 0$. Now, suppose we a priori fix some arbitrary strictly decreasing continuous function $\eta(v)$, with $\eta(0) = 1$ and $\eta(1) = 0$. Fix an integer $K \geq 1$. Consider the following jump probability rule (which uniquely determines the corresponding jump probability function $\eta_n^{(K)}(\cdot)$): a particle which gets an urge to jump chooses K other particles uniformly at random and actually jumps with (random) probability $\eta(k/K)$, where $k \geq 0$ is the (random) number of chosen particles that happen to be behind our particle. It is easy to see that, as $n \to \infty$,

$$\eta_n^{(K)}(v) \to \eta^{(K)}(v) \doteq \sum_{k=0}^{K} \frac{K!}{k!(K-k)!} v^k (1-v)^{K-k} \eta(k/K).$$

The results of [6] apply to such a system, with its mean-field limit $f(x, t)$ having $\eta^{(K)}(\cdot)$ in place of $\eta(\cdot)$. Note that, if K is large (but fixed!), $\eta^{(K)}(\cdot)$ can arbitrarily closely approximate $\eta(\cdot)$. To summarize, the results of [6] do *not* apply to show the convergence to a mean-field model with *any* a priori fixed jump probability function $\eta(\cdot)$. (Recall that in this paper we define a mean-field model formally, for *any* such function $\eta(v)$.) However, there always exists a function $\eta^{(K)}(\cdot)$, arbitrarily close to $\eta(\cdot)$, such that the results of [6] apply to show the convergence to a mean-field model with $\eta^{(K)}(\cdot)$ in place of $\eta(\cdot)$.

The second additional assumption in [6] is that, for any n, the system's initial state is such that the locations of the n particles are drawn *independently* from a given *absolutely continuous* distribution $f(\cdot, 0)$ with finite second moment.

Finally, [6] assumes that the jump size distribution $J(\cdot)$ is exponential. The results of [6] easily generalize to an arbitrary absolutely continuous distribution $J(\cdot)$.

3.3.2. *Detailed discussion of* [7]. The paper [7] considers a formally defined mean-field model (in the terminology of this paper), under the additional assumptions that $J(\cdot)$ is exponential, $J(y) = 1 - e^{-y}$, and $\eta(v) = (1-v)^K$.

The main result of [7] is basically as follows. Consider two mean-field models, $f_1(x, t)$ and $f_2(x, t)$, with well-defined, equal means, $\bar{f}_1(t) = \bar{f}_2(t) = \bar{f}_1(0) + vt = \bar{f}_2(0) + vt$. Then the derivative of the L_1-distance is negative, $(d/dt)\|f_1(\cdot, t) - f_2(\cdot, t)\|_1 < 0$, and is bounded away from 0 as long as $\|f_1(\cdot, t) - f_2(\cdot, t)\|_1$ is bounded away from 0. This, in turn, leads to the second result: *if a traveling wave shape ϕ exists*, then it is unique (up to a shift) and the convergence (3.5) to the traveling wave shape holds. These results, along with their proofs, extend to the more general case of arbitrary continuous strictly decreasing $\eta(v)$ and distribution $J(\cdot)$ satisfying (3.4).

As far as the existence of a traveling wave shape is concerned, in [7] it is demonstrated *only for the special case of exponential $J(\cdot)$ and $\eta(v) = (1 - v)^K$*, in which case (3.8) can be explicitly solved as follows. The derivation is not given in [7], so, for completeness, we give it here.

In the case $J(y) = 1 - e^{-y}$, (3.8) becomes an ordinary differential equation. Indeed, we have

$$v\phi'(x) = \int_{-\infty}^{x} \phi'(y)\eta(\phi(y))e^{-(x-y)}dy,$$

$$v\phi'(x)e^x = \int_{-\infty}^{x} \phi'(y)\eta(\phi(y))e^y dy,$$

$$v[\phi''(x) + \phi'(x)] = \phi'(x)\eta(\phi(x)). \tag{3.9}$$

Under the further assumption that $\eta(v) = (1 - v)^K$, (3.9) is easy to solve explicitly. In this case $v = 1/(K + 1)$. Define $\alpha(x) = 1 - \phi(x)$. Then

$$v[-\alpha''(x) - \alpha'(x)] = -\alpha'(x)\eta(1 - \alpha(x)) = -\alpha'(x)(\alpha(x))^K = -\frac{\left[(\alpha(x))^{K+1}\right]'}{K + 1},$$

$$-\alpha''(x) - \alpha'(x) = -\left[(\alpha(x))^{K+1}\right]',$$

$$-\alpha'(x) - \alpha(x) = -(\alpha(x))^{K+1} + c_1.$$

Recalling that $1 - \alpha(x)$ is a probability distribution function, we see that $c_1 = 0$ must hold, so

$$-\alpha'(x) - \alpha(x) = -(\alpha(x))^{K+1},$$

$$-\frac{(\log \alpha(x))'}{1 - (\alpha(x))^K} = 1,$$

$$\frac{1}{K}\left[\log\left[(\alpha(x))^{-K} - 1\right]\right]' = 1.$$

Solving for $\alpha(x)$, we finally obtain

$$\phi(x) = 1 - \left[1 + e^{K(x-c)}\right]^{-1/K}, \tag{3.10}$$

where the real number c is a (shift) parameter.

3.3.3. *Other related work.* Subsequently to [6, 7], there has been a line of work generally focused on modeling mean-field interaction between elements (particles) of a large system; cf. [1, 10–15]. The models and results in this line of work differ from those of this paper in that they consider different particle jump rules and/or different asymptotic regimes. But the

general common feature is that they study systems with mean-field ('global') interactions, when each particle's instantaneous behavior depends on the current distribution of the states of all the particles, as opposed to 'local' interactions, when a particle's instantaneous behavior depends only on the current states of its 'neighbors'. This naturally leads to the corresponding mean-field models describing the behavior of large-scale systems.

In particular, the paper [1] considers a related particle system with the jump probability depending not on a particle quantile within the empirical distribution $f^n(\cdot, t)$, but rather on the particle displacement with respect to the empirical distribution mean. For this system, under some technical assumptions, the paper proves the convergence of $f^n(x, t)$ to the corresponding mean-field model $f(x, t)$ (which is different from ours). This parallels the main result of [6] for our particle system. The paper also shows the existence of traveling wave shapes for that mean-field model, but only for the special case when the jump size distribution is exponential.

Further references and discussion of related work can be found in e.g. [1, 14].

4. Characterization of a traveling wave shape

Consider the following operator A, mapping a proper probability distribution $\psi(\cdot)$ on $\mathbb{R}$ into another distribution $A\psi(\cdot)$. Consider one particle, whose location $X(t)$ evolves as follows. The particle moves left at the constant speed v and sometimes jumps right, with i.i.d. jump sizes given by the distribution $J(\cdot)$. The particle gets urges to jump as a Poisson process of rate 1, and actually jumps with probability $\bar{\eta}(X(t), \psi)$, depending on the current particle location $X(t)$ and distribution the ψ. (Since the set of discontinuity points of ψ is at most countable, w.p.1 the process $X(\cdot)$ will never jump from points of discontinuity of ψ. So, equivalently, we can say that the jump probability is $\eta(\psi(X(t)))$.) The distribution $A\psi$ is defined as the stationary distribution of the process $X(\cdot)$. The following lemma shows that this stationary distribution exists and is unique.

Lemma 4.1. *Consider the single-particle movement process $X(\cdot)$ defined just above. This process is positive recurrent, and therefore has unique stationary distribution. Moreover, in steady state*

$$\mathbb{E}|X(t)|^{\ell-1} < \infty, \tag{4.1}$$

for the $\ell \geq 2$ from the condition (1.1)*.*

Proof. The distribution ψ is proper. Then, when $X(t) < 0$ with large absolute value, the process drift is positive, close to $1 - v$; when $X(t)$ is large positive, the process drift is negative, close to $-v$. The formal proof of positive recurrence is a straightforward application of the fluid limit technique [2, 3, 16, 17]. Indeed, it is easy to see that any compact subset of $\mathbb{R}$ is petite; see the definition in e.g. [2, 3]. (For a compact set, consider an interval $[-C, C]$ containing it. Then the sampling distribution can be chosen as, for example, the uniform distribution in $[0, 3C/v]$. Here we use the fact that in the interval $[0, 3C/v]$, with probability $1 - e^{-3C/v}$, the particle does not have an urge to jump and will simply move left at the constant speed v, covering the distance $3C$ and in particular the entire interval $[-2C, -C]$.) Furthermore, if we consider a sequence of processes $X^{(r)}(\cdot)$ with $|X^{(r)}(0)| = r$ and $r \to \infty$, and with $X^{(r)}(0)/r \to x(0) \in \{-1, 1\}$, then in our case it is easy to see that

$$X^{(r)}(rt)/r \xrightarrow{u.o.c.} x(t) \qquad \text{w.p.1}, \tag{4.2}$$

where $x(\cdot)$ is the deterministic Lipschitz function such that $x'(t) = 1 - v > 0$ when $x(t) < 0$, $x'(t) = -v < 0$ when $x(t) > 0$, and therefore $x(t)$ stays at 0 after it 'hits' 0. This implies positive recurrence.

Moreover, since the jump size distribution has finite ℓth moment, $\ell \geq 2$, using the results of [4], we obtain (4.1). Since, formally, our model is not within the framework of [4], we provide some details here. In our model the process $X(\cdot)$, with values in $\mathbb{R}$, corresponds to the queue length process $Q(\cdot)$ in [4], with values of $Q(t)$ being finite-dimensional vectors with non-negative integer components. In our case the process state norm is $|X(t)|$. The property (4.2), in particular, implies the existence of $T > 0$ such that

$$\frac{1}{r}|X^{(r)}(rT)| \xrightarrow{P} 0. \tag{4.3}$$

For our model, the analogue of Proposition 5.1 in [4] has the following form. There exists $T > 0$ such that

$$\lim_{r \to \infty} \frac{1}{r^\ell} \mathbb{E}|X^{(r)}(rT)|^\ell = 0, \tag{4.4}$$

where $X^{(r)}(\cdot)$ is a sequence of processes defined above in this proof. Given that we have (4.3), to prove (4.4) it suffices to show that the family of random variables

$$\frac{1}{r^\ell}|X^{(r)}(rT)|^\ell$$

is uniformly integrable. Denote by U_i, $i = 1, 2, \ldots$, the sequence of i.i.d. random variables, where U_i is equal in distribution to a jump size; we have $\mathbb{E}U_i^\ell < \infty$ by (1.1). Then the value of $|X^{(r)}(rT)|$ is stochastically dominated by

$$r + vrT + \sum_{i=1}^{\Pi(rT)} U_i,$$

where $\Pi(t)$ is Poisson process with constant rate 1. We see that, to show the uniform integrability of $|X^{(r)}(rT)|^\ell/r^\ell$, it suffices to show that the family

$$Z(t) = \frac{\left[\sum_{s=1}^{\Pi(t)} U_i\right]^\ell}{t^\ell}$$

is uniformly integrable for any positive increasing sequence $t \uparrow \infty$; this is true because $Z(t) \geq 0$ and, as $t \to \infty$, we have both $Z(t) \xrightarrow{P} [\mathbb{E}U_1]^\ell$ and $\mathbb{E}Z(t) \to [\mathbb{E}U_1]^\ell$. This proves (4.4). Since we have (4.4), which is the analogue of Proposition 5.1 in [4], the rest of the proof of Theorem 5.5 (and then Theorem 4.1(i)) in [4] applies as is. This, for our model, proves (4.1). $\square$

From now on in the paper, when we consider a distribution ϕ and the corresponding distribution $A\phi$, we will denote by $X(\cdot)$ the corresponding single-particle movement process which defines $A\phi$.

Lemma 4.2. *Consider any distribution ϕ and the corresponding process $X(\cdot)$ defining $A\phi$. Then the distribution $A\phi$ is $1/v$-Lipschitz.*

Proof. Recall that $\gamma = A\phi$ is nothing else but the stationary distribution of the single-particle movement process $X(\cdot)$ corresponding to ϕ. Consider the stationary regime of this process. Consider any x and any $\Delta > 0$. For any fixed initial position $X(0)$ of the particle and any $t > 0$, the expected total time in $[0, t]$ that the particle spends in the interval $(x, x + \Delta]$ is bounded above by

$$(t + 1)\Delta/v,$$

where $(t+1)$ is the expected number of time intervals *between jumps* the particle will have in $[0, t]$, and Δ/v is the maximum time the particle can possibly spend in $(x, x+\Delta]$ during each of those time intervals. Therefore, considering the stationary regime of the process,

$$\gamma(x+\Delta) - \gamma(x) \leq \frac{[t+1]\Delta/v}{t}.$$

This is true for any $t > 0$, which implies $\gamma(x+\Delta) - \gamma(x) \leq \Delta/v$, i.e. the $1/v$-Lipschitz continuity of $\gamma(\cdot)$. □

Lemma 4.3. *A probability distribution function $\phi(\cdot)$ is a traveling wave shape if $\phi = A\phi$.*

It can be shown that the converse is also true, and therefore a probability distribution function $\phi(\cdot)$ is a traveling wave shape *if and only if* $\phi = A\phi$. We do not need the 'only if' part for the proof of our main results.

Proof of Lemma 4.3. Suppose $\phi = A\phi$. Then, by Lemma 4.2, ϕ is Lipschitz. Consider any point x where the proper derivative $\phi'(x)$ exists. We need to show that (3.8) holds. Consider the single-particle movement process $X(\cdot)$ defining $A\phi$. In the stationary regime, the average rate at which the particle crosses the point x from right to left ('leaves' (x, ∞), to be precise) is $v\phi'(x)$; indeed, it is easy to see that the expected number of such crossings in the interval $[t, t+\Delta t]$ is equal, up to addition of $o(\Delta t)$ terms, to the probability that at time t the particle is in $(x, x+w\Delta t]$ and it does not jump in $[t, t+\Delta t]$—the latter probability itself is equal to $\hat{\gamma}'(x)w\Delta t + o(\Delta t)$. The average rate, in the stationary regime, at which the particle crosses from left to right ('enters' (x, ∞), to be precise) is the right-hand side of (3.8); indeed, such crossings may occur only when the particle jumps, and the right-hand side of (3.8) is the steady-state probability that an urge to jump results in a jump over the point x. These right-to-left and left-to-right crossing rates must be equal, thus proving (3.8). □

5. Traveling wave shapes within a finite frame

Let parameters $w \in (0, \infty)$ and $B_L, B_R \in (0, \infty)$ be fixed. We will now define the operator $A^{(w;B_L,B_R)}$ for a 'finite-frame' system; this operator is an analogue of the operator A for the original system. Fixed points of the operator $A^{(w;B_L,B_R)}$ will in turn define traveling wave shapes within a finite frame.

5.1. Definition and properties of the operator $A^{(w;B_L,B_R)}$

Consider a probability distribution function $\gamma = \gamma(\cdot)$ on $\mathbb{R}$, which is actually concentrated on the interval ('finite frame') $[-B_L, B_R]$; that is, $\gamma(-B_L-) = 0$, $\gamma(B_R) = 1$. Consider a single particle, whose location $X(t)$ evolves within the interval $[-B_L, B_R]$ as follows. The particle continuously moves left and sometimes makes jumps right. Between jumps, the particle moves left at the constant speed w, unless/until it reaches the left boundary $-B_L$, in which case the particle stops at $-B_L$ and stays there until the next jump—this is the 'regulation' at the left boundary. The particle gets urges to jump as a Poisson process of rate 1, and actually jumps with probability $\bar{\eta}(X(t), \gamma)$, depending on the current particle location $X(t)$ and distribution γ. (Here we *cannot* replace $\bar{\eta}(X(t), \gamma)$ with $\eta(\gamma(X(t)))$, because the process $X(\cdot)$ may spend non-zero time at the left boundary $-B_L$, where $\gamma(\cdot)$ may have a discontinuity.) The jump sizes are i.i.d., given by the distribution $J(\cdot)$; however, if a jump were to take the particle to the right of B_R, the particle lands at B_R instead—this is the 'regulation' at the right boundary. Then, the distribution $A^{(w;B_L,B_R)}\gamma$ is defined as the stationary distribution of the process $X(\cdot)$. It is easy to

see that the process $X(\cdot)$ is positive Harris recurrent: the times when the particle hits $-B_L$ are renewal points and the mean inter-renewal times are clearly finite. Therefore, $X(\cdot)$ indeed has unique stationary distribution. Clearly, $A^{(w;B_L,B_R)}\gamma$ is also concentrated on $[-B_L, B_R]$.

Lemma 5.1. *If $\hat{\gamma} = A^{(w;B_L,B_R)}\gamma$, then $\hat{\gamma}$ satisfies the following conditions:*

$$\hat{\gamma}(\cdot) \text{ is } 1/w - \text{Lipschitz in } [-B_L, B_R]; \tag{5.1}$$

$$\hat{\gamma}(-B_L-) = 0 < \hat{\gamma}(-B_L), \qquad \hat{\gamma}(B_R) = 1; \tag{5.2}$$

almost all points $-B_L < x < B_R$ (with respect to Lebesgue measure) are regular, *in that a proper derivative $\hat{\gamma}'(x)$ exists, and at each regular point*

$$w\hat{\gamma}'(x) = \hat{\zeta}(x) \doteq \int_{-B_L-}^{x} d\hat{\gamma}(y)\bar{\eta}(y;\gamma)\bar{J}(x-y); \tag{5.3}$$

the function $\hat{\zeta}(x)$ is RCLL in $[-B_L, B_R]$, and $\hat{\zeta}(x-) \geq \hat{\zeta}(x)$, $\tag{5.4}$

and

$$w\frac{d^+}{dx}\hat{\gamma}(x) = \hat{\zeta}(x), \forall x \in [-B_L, B_R), \qquad w\frac{d^-}{dx}\hat{\gamma}(x) = \hat{\zeta}(x-), \forall x \in (-B_L, B_R]. \tag{5.5}$$

Proof. Recall that $\hat{\gamma}$ is nothing else but the stationary distribution of the location process $X(\cdot)$ of a single particle, corresponding to γ. Consider the stationary regime of this process. Then the $1/w$-Lipschitz property of $\hat{\gamma}$ in $(-B_L, B_R)$ is proved in exactly same way as in the proof of Lemma 4.2, with w replacing v. Given that $\hat{\gamma}$ is right-continuous, it is $1/w$-Lipschitz in $[-B_L, B_R]$ as well. The properties (5.2) are obvious from the structure of the particle movement process $X(\cdot)$: clearly, in steady state the process spends zero time exactly at the right boundary point B_R and a non-zero fraction of time at the left boundary point $-B_L$.

Since $\hat{\gamma}(\cdot)$ is Lipschitz in $(-B_L, B_R)$, it is absolutely continuous, and then almost all points $-B_L < x < B_R$ are regular (that is, the proper derivative $\hat{\gamma}'(x)$ exists). Consider any regular point $-B_L < x < B_R$. To prove (5.3), essentially the same argument as in the proof of Lemma 4 applies. Indeed, in the stationary regime, the average rate at which the particle crosses the point x from right to left ('leaves' $(x, B_R]$, to be precise) is $w\hat{\gamma}'(x)$ (by the same argument as in the proof of Lemma 4.3). The average rate, in the stationary regime, at which the particle crosses from left to right ('enters' $(x, B_R]$, to be precise) is $\hat{\zeta}(x)$; indeed, such crossings may occur only when the particle jumps, and $\hat{\zeta}(x)$ is the steady-state probability that an urge to jump results in a jump over the point x. These right-to-left and left-to-right crossing rates must be equal, which gives (5.3).

The property (5.4) easily follows from a direct analysis of the expression for $\hat{\zeta}(x)$, using the facts that $\bar{J}(\cdot)$ is non-increasing and RCLL, and that the measure given by $\hat{\gamma}$ has exactly one atom at $-B_L$ and is absolutely continuous in $(-B_L, B_R)$. We omit details. Finally, (5.5) follows from (5.4) and (5.3). □

Let $\Gamma(w; B_L, B_R)$ denote the set of distribution functions γ concentrated on the interval $[-B_L, B_R]$, $B_L, B_R > 0$, which are $1/w$-Lipschitz on $[-B_L, B_R]$; here $w > 0$, $B_L > 0$, and $B_R > 0$ are parameters 'attached to' γ. We define $\Gamma = \cup_{w,B_L,B_R>0}\Gamma(w; B_L, B_R)$. In particular, $\gamma \in \Gamma$ implies that $\gamma(B_R-) = \gamma(B_R) = 1$ and $\gamma(B_L-) = 0 \leq \gamma(B_L)$. So, γ may have a single atom, at $-B_L$. We endow Γ with the following natural topology: the convergence $\gamma^{(k)} \xrightarrow{\Gamma} \gamma$ is

defined as $w^{(k)} \to w$, $B_L^{(k)} \to B_L$, $B_R^{(k)} \to B_R$, and $\gamma^{(k)}(y) \to \gamma(y)$ for all $y \in (-B_L, B_R)$. In particular, $\gamma^{(k)} \overset{\Gamma}{\to} \gamma$ implies $\gamma^{(k)} \overset{J_1}{\to} \gamma$. In the special case when $w^{(k)} \equiv w$, $B_L^{(k)} \equiv B_L$, $B_R^{(k)} \equiv B_R$, convergence $\gamma^{(k)} \overset{\Gamma}{\to} \gamma$ is equivalent to $\gamma^{(k)} \overset{u}{\to} \gamma$.

Lemma 5.2. *The map $A^{(w;B_L,B_R)}\gamma$, with $(w; B_L, B_R)$ being the parameters of γ, is continuous in $\gamma \in \Gamma$ (in the Γ-topology).*

Proof. As $\gamma^{(k)} \overset{\Gamma}{\to} \gamma$, denote the corresponding single-particle processes by $X_k(\cdot)$ and $X(\cdot)$. Clearly, the sequence $A^{(w^{(k)};B_L^{(k)},B_R^{(k)})}\gamma^{(k)}$ is tight. Consider any subsequential limit of it, $\hat{\gamma}$, along a subsequence of k. Consider each of the processes $X_k(\cdot)$ and $X(\cdot)$ on a finite time interval $[0, T]$. These processes can be naturally coupled so that the following property holds w.p.1 along the chosen subsequence: if either (a) $X(0) = -B_L$ and $X_k(0) = -B_L^{(k)}$ for all k, or (b) $X_k(0) \to X(0) > -B_L$, then

$$(X_k(t), 0 \le t \le T) \overset{J_1}{\to} (X(t), 0 \le t \le T).$$

Given this kind of continuity, it is straightforward to see that the limit $\hat{\gamma}(\cdot)$ of the stationary distributions of $X_k(\cdot)$ must be the stationary distribution $A^{(w;B_L,B_R)}\gamma$ of $X(\cdot)$. $\square$

5.2. Definition and properties of finite-frame traveling wave shapes

Denote by $\mathcal{T}(w; B_L, B_R)$ the set of those distributions γ which are fixed points of $A^{(w;B_L,B_R)}$, i.e. which satisfy $\gamma = A^{(w;B_L,B_R)}\gamma$. Any $\gamma \in \mathcal{T}(w; B_L, B_R)$ we will call *a traveling wave shape for the finite frame $[-B_L, B_R]$ and speed w*. If such a traveling wave shape γ exists and is unique, i.e. the cardinality of $\mathcal{T}(w; B_L, B_R)$ is 1, we will slightly abuse notation by writing $\gamma = \mathcal{T}(w; B_L, B_R)$.

Lemma 5.3. *Let w, B_L, $B_R > 0$ be fixed. Then the following hold:*

(i) $\gamma = \mathcal{T}(w; B_L, B_R)$ exists and is unique.

(ii) $\gamma = \mathcal{T}(w; B_L, B_R)$ if and only if it satisfies the following conditions:

$$\gamma(\cdot) \text{ is } 1/w - \text{Lipschitz in } [-B_L, B_R]; \tag{5.6}$$

$$\gamma(-B_L-) = 0 < \gamma(-B_L), \qquad \gamma(B_R) = 1; \tag{5.7}$$

almost all points $-B_L < x < B_R$ (with respect to Lebesgue measure) are regular, *in that a proper derivative $\gamma'(x)$ exists, and at each regular point*

$$w\gamma'(x) = \zeta(x) \doteq \int_{-B_L-}^{x} d\gamma(y)\bar{\eta}(y; \gamma)\bar{J}(x-y); \tag{5.8}$$

the function $\zeta(x)$ is RCLL in $[-B_L, B_R]$, and $\zeta(x-) \ge \zeta(x)$, \tag{5.9}

and

$$w\frac{d^+}{dx}\gamma(x) = \zeta(x), \forall x \in [-B_L, B_R), \qquad w\frac{d^-}{dx}\gamma(x) = \zeta(x-), \forall x \in (-B_L, B_R]; \tag{5.10}$$

$$\inf_{[-B_L,B_R]} \zeta(x) > 0. \tag{5.11}$$

Proof of Lemma 5.3. The proof consists of the following three claims.   □

Claim 1. $\gamma \in \mathcal{T}(w; B_L, B_R)$ exists.

Claim 2. If $\gamma \in \mathcal{T}(w; B_L, B_R)$, then it satisfies the conditions (5.6)–(5.11).

Claim 3. A distribution γ concentrated on $[-B_L, B_R]$ and satisfying (5.6)–(5.11) is unique.

Proof of Claim 1. The existence follows from the Brouwer fixed point theorem. (Cf. [8, Theorem XVI.5.1] for its more general form—Kakutani's theorem.) Indeed, by Lemma 5.2, for a fixed set of positive parameters $w, B_L, B_R, A^{(w;B_L,B_R)}\gamma$ continuously (in the uniform convergence topology) maps $\Gamma(w; B_L, B_R)$ into itself. The set $\Gamma(w; B_L, B_R)$ is convex and compact. Therefore, a fixed point $\gamma = A^{(w;B_L,B_R)}\gamma$ exists.   □

Proof of Claim 2. The properties (5.6)–(5.10) follow from Lemma 5.1 and the properties (5.1)–(5.5).

Next, observe that $\gamma(x)$ must be strictly increasing in $[-B_L, B_R]$. If not, we could find an interval $[y_1, y_2] \subset [-B_L, B_R)$ such that $y_1 < y_2$, $\gamma(y_2) = \gamma(y_1)$ (and then $(d^+/dx)\gamma(y_1) = 0$), and $\gamma(x) < \gamma(y_1)$ for $x < y_1$; but, by (5.10) and (5.8), $(d^+/dx)\gamma(y_1) = \zeta(y_1)/w > 0$, a contradiction.

Let us now prove (5.11). First, from (5.8) and the fact that $\gamma(x)$ is strictly increasing in $[-B_L, B_R]$, observe that $\liminf_{x \uparrow B_R} \zeta(x) > 0$. Therefore, if $\inf_{[-B_L, B_R]} \zeta(x) = 0$ were to hold, we would have a point $y \in [-B_L, B_R)$ such that $\zeta(y) = 0$. (Here we use (5.9).) But $\gamma(x)$ is strictly increasing in a neighborhood of y, which, again by (5.8), implies $\zeta(y) > 0$. This completes the proof of (5.11) and of Claim 2.   □

Proof of Claim 3. First, note that any γ satisfying (5.6)–(5.11) is such that the following holds: the inverse $\gamma^{-1}(v)$ for $v \in (0, 1]$ is continuous and Lipschitz, and is strictly increasing in $[\gamma(-B_L), 1]$. Moreover, the mapping $v = \gamma(x)$ gives a one-to-one correspondence between regular points of γ in $[-B_L, B_R]$ and regular points of γ^{-1} in $[\gamma(-B_R), 1]$, with the derivatives at corresponding regular points satisfying $(d/dv)\gamma^{-1}(v) = 1/\gamma'(x)$. Furthermore, the latter relation holds for the right derivative for any $x \in [-B_L, B_R)$ and corresponding $v = \gamma(x)$, and for the left derivative for any $x \in (-B_L, B_R]$.

The proof of uniqueness is by contradiction. Suppose γ_1 and γ_2 are two different distributions satisfying the conditions (5.6)–(5.11). Let us define $\alpha_i(x) \doteq w(d^+/dx)\gamma_i(x)$, $i = 1, 2$; then, by the first equation in (5.10), $\alpha_i(x) = \zeta_i(x)$.

Suppose, for example, that $\gamma_1(y) > \gamma_2(y)$ for at least one $y \geq 0$. Then let us call

$$h = \max_v \left[\gamma_2^{-1}(v) - \gamma_1^{-1}(v) \right] > 0$$

the 'horizontal distance' between γ_1 and γ_2. It is well-defined, by the properties of γ_1 and γ_2. Let v_* be the minimum v at which the horizontal distance is attained. Clearly, $0 < v_* < 1$. Let $y_1 = \gamma_1^{-1}(v_*) < y_2 = \gamma_2^{-1}(v_*)$. Then we must have $(d^+/dx)\gamma_1(y_1) \leq (d^+/dx)\gamma_2(y_2)$ or, equivalently, $\alpha_1(y_1) \leq \alpha_2(y_2)$. Given the choice of v_*, $y_1 - \gamma_1^{-1}(v) \leq y_2 - \gamma_2^{-1}(v)$ for all $v \in [0, 1]$, and then $\zeta_1(y_1) \geq \zeta_2(y_2)$. Since $\alpha_i(y_i) = \zeta_i(y_i)$, we must have

$$\alpha_1(y_1) = \alpha_2(y_2) = \zeta_1(y_1) = \zeta_2(y_2).$$

Note that if, for example, $\bar{J}(y) < 1$ for all $y > 0$, then we immediately obtain a contradiction, because in this case $\zeta_1(y_1) > \zeta_2(y_2)$. The case of general jump size distribution J requires some more details, which are as follows.

We must have that $\bar{J}(y_2-) = 1$, i.e. the jump size is at least y_2. Indeed, by the choice of ν_*, $y_1 - \gamma_1^{-1}(\nu) < y_2 - \gamma_2^{-1}(\nu)$ for all $\nu < \nu_*$. Also, recall that both inverse functions $\gamma_i^{-1}(\nu)$ are continuous. This means that as we increase ν in the interval $[0, \nu_*)$, the interval $\left[y_1 - \gamma_1^{-1}(\nu), y_2 - \gamma_2^{-1}(\nu)\right]$ continuously changes ('moves left') from $[y_1 + B_L, y_2 + B_L]$ to (in the limit!) $[0, 0]$, while remaining of non-zero length. If at least one of points $y \in [0, y_2)$ were a point of decrease of $\bar{J}$, we would have $\bar{J}\left(y_1 - \gamma_1^{-1}(\nu)\right) > \bar{J}\left(y_2 - \gamma_2^{-1}(\nu)\right)$ on a subset of $[0, \nu_*)$ having non-zero Lebesgue measure, which would imply $\zeta_1(y_1) > \zeta_2(y_2)$, contradicting the definition of ν_*.

Consider now

$$\nu^* = \max\left\{\bar{\nu} \geq \nu_* \,|\, \gamma_2^{-1}(\nu) - \gamma_1^{-1}(\nu) = h, \, \forall \nu \in [\nu_*, \bar{\nu}]\right\}.$$

We must have $\nu^* < 1$. Let $y_i^* = \gamma_i^{-1}(\nu^*)$. It is easy to see that $y_2^* = \max\left\{y \geq y_2 \,|\, \bar{J}(y-) = 1\right\}$, i.e. y_2^* must be exactly the smallest point of decrease of $\bar{J}$. Indeed, $\bar{J}$ cannot have a point of decrease $\hat{v} \in [\nu_*, \nu^*)$. If such a point $\hat{v}$ existed, then using an argument analogous to that we used just above to show that $\bar{J}(y_2-) = 1$, we would obtain that $\gamma_1^{-1}(\nu) - \gamma_2^{-1}(\nu)$ would be strictly increasing in a small interval immediately to the right of the point $\hat{v}$, which would contradict the definition of the horizontal distance h. On the other hand, if y_2^* were *not* a point of decrease of $\bar{J}$, i.e. $\bar{J}(\hat{v}) = 1$ for some $\hat{y} > y_2^*$, then, again using essentially the same argument, we would obtain that $\gamma_1^{-1}(\nu) - \gamma_2^{-1}(\nu)$ would have to remain constant in a small interval immediately to the right of the point ν^*, which would contradict the definition of ν^*.

Finally, if y_2^* is a point of decrease of $\bar{J}$, then, using essentially the same argument once again, we obtain that $\gamma_1^{-1}(\nu) - \gamma_2^{-1}(\nu)$ must be strictly increasing in a small interval immediately to the right of the point ν^*, which contradicts the definition of the horizontal distance h. $\qquad \square$

The proof of Lemma 5.3 is complete.

Lemma 5.4. (Continuity, monotonicity, and shift properties.)

(i) $\gamma = \mathcal{T}(w; B_L, B_R)$, as an element of Γ, is continuous in $(w; B_L, B_R)$.

(ii) $\gamma = \mathcal{T}(w; B_L, B_R)$, as a probability distribution, is monotone (in the sense of $\preceq$) in each of the parameters w, B_L, and B_R, given that the other two parameters are fixed. Specifically, γ is monotone non-increasing in w and B_L, and monotone non-decreasing in B_R.

(iii) Suppose B_L and B_R are fixed. If $w \uparrow \infty$ [resp., $w \downarrow 0$], then $\gamma = \mathcal{T}(w; B_L, B_R)$, as a probability distribution, weakly converges to the Dirac distribution concentrated at $-B_L$ [resp., at B_R]. Consequently, for any fixed B_L, B_R, any fixed $y \in (-B_L, B_R)$, and any fixed $\nu \in (0, 1)$, there exists $w > 0$ such that $\gamma^{-1}(\nu) = y$. In other words, by changing w we can always move a given quantile of γ to any point within $(-B_L, B_R)$.

(iv) For any fixed parameters w, B_L, and B_R, and any $c \in \mathbb{R}$, $\gamma = \mathcal{T}(w; B_L, B_R)$ implies $\theta_c \gamma = \mathcal{T}(w; B_L - c, B_R + c)$. In other words, a shift of the finite frame by c results in the same shift of the corresponding traveling wave shape.

Proof. (i) The continuity follows immediately from the continuity of the operator $A^{(w; B_L, B_R)}$ on Γ (Lemma 5.2) and the uniqueness of $\gamma = \mathcal{T}(w; B_L, B_R)$ (Lemma 5.3(i)).

(ii) The monotonicity is proved by contradiction, using the horizontal distance, in almost exactly the same way as in the proof of uniqueness (Claim 3) in Lemma 5.3.

(iii) The case $w \uparrow \infty$ is obvious, because γ is $1/w$-Lipschitz in $[B_L, B_R]$ and $\gamma(B_R) = 1$. Suppose $w \downarrow 0$. Then $\gamma = \mathcal{T}(w; B_L, B_R)$ is monotone non-increasing in w. It suffices to show that, for each $y < B_R$, $\lim \gamma(y) = 0$. Indeed, fix $y < B_R$ and any $y_1 \in (y, B_R)$. Clearly, $\lim \gamma(y_1) < \nu_1 < 1$. Fix small $\epsilon > 0$, so that $y < y_1 - \epsilon < y_1$. Then, uniformly in all sufficiently small $w > 0$, the corresponding particle process $X(\cdot)$ is such that when it is to the left of the point y_1, it will jump with rate at least $\eta(\nu_1)$, while it moves left at the small speed w; this in turn easily implies that the steady-state probability of the particle being to the left of $y_1 - \epsilon$ vanishes.

(iv) This is obvious from the definition of a finite-frame traveling wave shape. $\qquad\square$

6. Proofs of the main results, Theorems 3.1 and 3.2

6.1. Proof of Theorem 3.1(i): existence

The high-level outline of the proof is as follows. We obtain a traveling wave shape ϕ as a limit of finite-frame traveling wave shapes γ_B, as $B \to \infty$, for the finite frame $[-B, B]$ and with the speed w_B chosen so that the median of γ_B is at 0. The first step, Lemma 6.1, shows that, necessarily, $w_B \to v$. The next step is to show, in Lemma 6.2, that the family of distributions $\{\gamma_B\}$ is tight. The proof of Lemma 6.2 involves considering the single-particle movement processes $X_B(\cdot)$ corresponding to γ_B, and their space-/time-rescaled versions $X_{(C)}(\cdot) = X_B(C \cdot)/C$, with $C = C(B) \to \infty$ as $B \to \infty$. (The supplementary Lemma 6.4, used in the proof of Lemma 6.2, is given in Section 6.3.) After the tightness of $\{\gamma_B\}$ is established, Lemma 6.3 completes the proof of Theorem 3.1(i) by demonstrating that any subsequential limit ϕ of γ_B must be a traveling wave shape. At the beginning of the proofs of Lemmas 6.1, 6.2, and 6.4 we give intuition and outlines specifically for those proofs.

We now proceed with the formal proof. Let us consider the finite-frame system with $B_L = B_R = B$, i.e. within the frame $[-B, B]$. Let $B \uparrow \infty$ along some fixed sequence. For each B, we will choose the speed w_B such that the corresponding γ, which we denote by γ_B, has its median exactly at 0, i.e. $\gamma_B(0) = 1/2$ or, equivalently, $\gamma_B^{-1}(1/2) = 0$. (It does not have to be the median; we could fix any $\nu_0 \in (0, 1)$ and choose the speed w_B such that $\gamma_B^{-1}(\nu_0) = 0$.) Such w_B and γ_B exist by Lemma 5.4(iii), and $0 < w_B < \infty$. The corresponding single-particle movement process $X(\cdot)$ we will denote by $X_B(\cdot)$.

Lemma 6.1. *Necessarily,* $\lim_{B \to \infty} w_B = v$.

Proof. The proof is by contradiction. The basic intuition is as follows. Suppose, for example, that $\limsup_{B \to \infty} w_B > v$. Then there exists a constant $w_\infty > v$ such that $w_B > w_\infty > v$ along some subsequence of B. Then we show that the distributions γ_B would have to stay tight around the left end $-B$ of the frame. This is, informally speaking, due to the fact that, if in steady state $X_B(\cdot)$ were to stay away from $-B$ most of the time, then its average drift would be bounded above by $v - w_\infty < 0$, while it must be 0 for any B. But if γ_B stays tight around $-B$, then as $B \to \infty$, the median of γ_B cannot stay at 0, which contradicts the γ_B definition. The contradiction in the case $\liminf_{B \to \infty} w_B < v$ is obtained similarly—we would have the distributions γ_B staying tight around the right end B of the frame.

We proceed with the formal proof. Suppose w_B does not converge to v. Consider, first, the case $\limsup_{B \to \infty} w_B > v$. Then we can choose a constant $w_\infty > v$ and a subsequence of B along which $w_B > w_\infty > v$. For all large B, consider $\gamma_B^{up} = \mathcal{T}(w_\infty; B, B)$ corresponding to the speed w_∞; also consider the corresponding single-particle process $X_B^{up}(\cdot)$. By monotonicity in w (Lemma 5.4(ii)), $\gamma_B \preceq \gamma_B^{up}$ for all large B. By monotonicity in the right end of the

frame (Lemma 5.4(ii)) and the shift property (Lemma 5.4(iv)), we observe that the distribution $\theta_B \gamma_B^{up}$ (i.e., the distribution γ_B^{up} shifted right by B, and therefore concentrated on $[0, 2B]$) will be monotone non-decreasing (in the sense of $\preceq$) in B. Then $\theta_B \gamma_B^{up}$ must uniformly converge, $\theta_B \gamma_B^{up} \xrightarrow{u} \gamma_\infty$, to some distribution γ_∞ concentrated on $[0, \infty)$, and the function γ_∞ is $1/w_\infty$-Lipschitz on $[0, \infty)$. (This follows from the fact that all functions $\theta_B \gamma_B^{up}$ are uniformly $1/w_\infty$-Lipschitz in $[0, \infty)$.) Moreover,

$$\gamma_\infty(0) > 0. \tag{6.1}$$

Indeed the particle process $X_B^{up}(\cdot)$ is such that the steady-state average drift of the particle is bounded above by $-\left(1 - \gamma_B^{up}(-B)\right)w_\infty + v$; if (6.1) did not hold, then, as $B \to \infty$, this upper bound would converge to $-\left(1 - \gamma_\infty(0)\right)w_\infty + v = -w_\infty + v < 0$, which is impossible.

To obtain a contradiction, it remains to show that the distribution γ_∞ is proper, i.e. $\lim_{y \to \infty} \gamma_\infty(y) = 1$. Indeed, this will imply that the median of γ_B^{up}, and then of γ_B as well, will eventually become less than 0, which contradicts the definition of γ_B.

Let us show that the distribution γ_∞ is proper. Suppose not; that is, $\lim_{y \to \infty} \gamma_\infty(y) = v_\infty^* < 1$. Let $v^* = \eta^{-1}(w_\infty)$.

The case $v_\infty^* > v^*$ is impossible. Indeed, in this case, pick a small $\epsilon > 0$ and any fixed $H > 0$ such that $\gamma_\infty(H) > v^* + \epsilon$. Then, uniformly in all large B, when the particle process $X_B^{up}(\cdot) + B$ ('living' in $[0, 2B]$) is to the right of H, the particle jumps right at most at the rate $\eta(v^* + \epsilon) < w_\infty$. We also have the following: when the process $X_B^{up}(\cdot) + B$ 'jumps over the point H', we have that uniformly in all B and on the jump starting point in $[0, H]$, the distribution of the 'overshoot-over-H' distance V is uniformly stochastically bounded above by some proper distribution with finite mean. (This is obvious if the jump size distribution $J(\cdot)$ has finite support. Otherwise, $\mathbb{P}\{V > y\} \leq \bar{J}(y)/\bar{J}(H)$.) Using these properties, we can construct a proper stochastic upper bound, uniform in B, on the stationary distribution of $X_B^{up}(\cdot) + B$, which is $\theta_B \gamma_B^{up}$. This contradicts the fact that the distribution γ_∞ is not proper.

Consider the case $v_\infty^* \leq v^*$. Consider the particle process $X_\infty(\cdot)$, taking values in $\mathbb{R}_+ = [0, \infty)$, that is defined in exactly the same way as the process defining the (finite-frame) operator value $A^{(w; B_L, B_R)}\gamma$, except that $X_\infty(\cdot)$ corresponds to the speed w_∞ and distribution γ_∞ on $\mathbb{R}_+ = [0, \infty)$ (which is not necessarily proper); the lower boundary is 0 (at which regulation does occur); and there is no upper boundary (and no regulation from above—no forward jump is ever 'truncated'). The process $X_\infty(\cdot)$ is regenerative, with regenerations occurring when the particle hits 0. Given that for any finite $y \geq 0$, $\eta(\gamma_\infty(y)) > w_\infty$, the process cannot possibly be positive recurrent. (Indeed, even if the rate of jumps is constant at w_∞, i.e. lower than it actually is, the process $X_\infty(\cdot)$ models the workload of a queue where the workload is depleted at the constant rate w_∞ and new workload arrives as a random process with average rate exactly w_∞. Therefore, $X_\infty(\cdot)$ is stochastically bounded below by a non-positive recurrent process.) Consider two sub-cases: (a) $X_\infty(\cdot)$ is transient (which is certainly the case if $v_\infty^* < v^*$, but logically also possible if $v_\infty^* = v^*$), and (b) $X_\infty(\cdot)$ is null-recurrent (which is possible only when $v_\infty^* = v^*$).

Observe that the processes $X_B^{up}(\cdot) + B$ for all B and the process $X_\infty(\cdot)$, all starting at 0, can be naturally coupled so that, as $B \to \infty$, the trajectory of $X_B^{up}(\cdot) + B$ converges (in the Skorokhod topology) to that of $X_\infty(\cdot)$ w.p.1.

Sub-case (a): If $X_\infty(\cdot)$ is transient, then for some fixed $\delta > 0$ and any $H > 0$, for all sufficiently large B, the probability that $X_B^{up}(\cdot) + B$ will up-cross level H in a regeneration cycle is at least δ. This implies that the mean duration of the regeneration cycle of $X_B^{up}(\cdot) + B$ has to go to ∞ as $B \to \infty$ (because the time to return to 0 from a point $y \geq H$ is at least $H/(w_\infty/2)$). This

implies that the steady-state probability $\theta_B \gamma_B^{up}(0)$ of $X_B^{up}(\cdot) + B$ being at 0 will vanish. Indeed, the mean time the process spends at 0 within one regeneration cycle is uniformly bounded above by an exponential random variable with mean $1/v$ (because the rate at which the particle jumps forward when it sits at the left boundary point is at least $\int_0^1 \eta(v)dv = v$). But vanishing $\theta_B \gamma_B^{up}(0)$ would mean $\gamma_\infty(0) = 0$, which contradicts (6.1).

Sub-case (b): If $X_\infty(\cdot)$ is null-recurrent, then the regeneration cycle of $X_\infty(\cdot)$ is finite w.p.1, and has infinite mean. Recall the natural coupling of $X_B^{up}(\cdot) + B$ and $X_\infty(\cdot)$. Given this coupling, for any fixed $H > 0$, on the event that $X_\infty(\cdot)$ does not up-cross H in a regeneration cycle, the regeneration cycle length of $X_B^{up}(\cdot) + B$ converges to that of $X_\infty(\cdot)$ almost surely. Using Fatou's lemma, we see that, as $B \to \infty$, $\liminf$ of the mean regeneration cycle duration of $X_B^{up}(\cdot) + B$ is at least that of $X_\infty(\cdot)$, which is infinity. This, again, leads to vanishing $\theta_B \gamma_B^{up}(0)$ and a contradiction with (6.1).

This completes the proof that $\limsup_{B\to\infty} w_B > v$ is impossible.

The contradiction to $\liminf_{B\to\infty} w_B < v$ is obtained similarly. In this case, we can construct lower bounds on γ_B. The condition (6.1) is replaced by the condition that the steady-state rate of the particle hitting the right boundary B converges down to a positive number. Regeneration time points are when the particle hits the right boundary B. We omit details. $\qquad\square$

Lemma 6.2. *The family of distributions $\{\gamma_B\}$ is tight.*

Proof. The proof is by contradiction. Its basic outline is as follows. If $\{\gamma_B\}$ is not tight, we can find a subsequence of B, and a corresponding sequence of scaling factors $C = C(B) \uparrow \infty$, such that the rescaled distributions $\psi_{(C)}(y) = \gamma_B(Cy)$ are such that, say, some v-quantile ($v > 1/2$) of $\psi_{(C)}$ remains at the point 1. Each $\psi_{(C)}$ is a finite-frame traveling wave shape; its frame is $[-B/C, B/C]$, the corresponding single-particle process is $X_{(C)}(t) = X_B(Ct)/C$, and the speed is still w_B. We show that the family of distributions $\{\psi_{(C)}\}$ is tight. This is done by considering a quadratic Lyapunov function for $X_{(C)}(t)$ and its steady-state drift, and showing that $\mathbb{E}|X_{(C)}(\infty)|$ remains uniformly bounded. Given the tightness of the sequence $\{\psi_{(C)}\}$, we consider a subsequence of it that converges to some proper distribution $\psi_{(\infty)}$. Using Lemma 6.4 we show that $\psi_{(\infty)}$ must be concentrated on at most two atoms, $z_1 \leq 0$ and $z_2 \geq 0$. The case of one atom, $z_1 = z_2$, is impossible because it leads to a contradiction with the v-quantile ($v > 1/2$) of $\psi_{(C)}$ remaining at the point 1. Then we must have either $z_1 < 0$ or $z_2 > 0$. In this case we again use a quadratic Lyapunov function for $X_{(C)}(t)$ to show that, for large C, its steady-state drift would be negative; roughly, this is because the probability of $X_{(C)}(\infty) \geq z_2$ is about $1 - v$ and the probability of $X_{(C)}(\infty) \leq z_1$, and the drift of the quadratic Lyapunov function is strictly negative when $X_{(C)}(\infty) \geq z_2 > 0$ and non-positive when $X_{(C)}(\infty) \leq z_1 \leq 0$.

We proceed with the formal proof by contradiction. Suppose $\{\gamma_B\}$ is not tight. Then there exists fixed $v > 1/2$ and a subsequence $B \uparrow \infty$ along which $C = C(B) = \max\{\gamma_B^{-1}(v), -\gamma_B^{-1}(1-v)\} \uparrow \infty$. Without loss of generality, we can choose v sufficiently close to 1 so that

$$\eta(1-v) > v > \eta(v). \qquad (6.2)$$

Note that, for each B, either $\gamma_B^{-1}(v) = C$ or $\gamma_B^{-1}(1-v) = -C$; also, $\gamma_B(C) - \gamma_B(-C) \geq v - (1-v) = 2v - 1$.

For each B (with the corresponding C), consider the particle process with space compressed by C and time sped up by C, i.e. the process $X_{(C)}(t) = X_B(Ct)/C$. For each B, the process lives within the frame $[-C_0, C_0]$, where $C_0 = B/C$. It is possible that $C_0 \to \infty$ as $B \to \infty$. Denote by $\psi_{(C)}(y) = \gamma_B(Cy)$ the corresponding scaled distribution. From now on in this proof, when we

say 'for a given B' or 'for a given C' we mean 'for a given pair (B, C) of B and C, corresponding to each other'.

Note that, for each C, either $\psi_{(C)}^{-1}(v) = 1$ or $\psi_{(C)}^{-1}(1 - v) = -1$; also, $\psi_{(C)}(1) - \psi_{(C)}(-1) \geq 2v - 1$.

Let us prove that

$$\text{the family of distributions } \{\gamma_C\} \text{ is tight.} \tag{6.3}$$

This is trivially true if C_0 remains bounded as $B \to \infty$. Therefore, it suffices to prove (6.3) for a subsequence of B such that $C_0 = B/C \to \infty$. Let us consider such a subsequence.

Let $g(y) = y^2/2$. Denote by $P_{(C)}^t(y, H)$ the transition function of the process $X_{(C)}(t)$. According to our definitions, for any fixed C, $\psi_{(C)} P_{(C)}^t g = \psi_{(C)} g$ for all $t \geq 0$; here we view the transition function $P_{(C)}^t$ and the distribution $\psi_{(C)}$ as operators (see Section 2).

Note that $\psi_{(C)} = \mathcal{T}(w_B; C_0, C_0)$, and that $X_{(C)}(\cdot)$ is the corresponding (finite-frame) particle process. Denote by $\hat{X}_{(C)}(\cdot)$ a single-particle process, corresponding (like $X_{(C)}(\cdot)$) to the distribution function $\psi_{(C)}$ and speed w_B, but different from $X_{(C)}(\cdot)$ in that it evolves in $(-\infty, \infty)$, i.e. no regulation keeping the process within the finite frame $[-C_0, C_0]$ is applied. The particle continues to move left at speed w_B even when it is at or to the left of $-C_0$, in which case its jump rate remains the same as it would be at the point $-C_0$, namely $\bar{\eta}(-C_0, \psi_{(C)})$. The particle also can 'jump over' the point C_0; when it is at or to the right of C_0, it moves left at speed w_B and its jump rate is as it would be at the point C_0, namely $\bar{\eta}(C_0, \psi_{(C)}) = \eta(\psi_{(C)}(C_0)) = \eta(1) = 0$. Denote by $\hat{P}_{(C)}^t(y, H)$ the transition function of the process $\hat{X}_{(C)}(t)$. We claim that, for any sufficiently large fixed C, uniformly in $y \in [-C_0, C_0]$,

$$\left[P_{(C)}^t g\right](y) - \left[\hat{P}_{(C)}^t g\right](y) \leq o(t), \tag{6.4}$$

where $o(t)$ is a positive function (which may depend on C) such that $o(t)/t \to 0$ as $t \to 0$. The proof of (6.4) is given in Section 6.2.

It is easy to check directly that, uniformly in $y \in [-C_0, C_0]$,

$$\lim_{t \downarrow 0} \frac{1}{t} \left[\hat{P}_{(C)}^t - I\right] g(y) = \hat{G}_{(C)} g(y) \doteq -w_B g'(y) + C\eta(\psi_{(C)}(y)) \int_0^\infty dJ(C\xi)[g(y + \xi) - g(y)],$$

and we have

$$\hat{G}_{(C)} g(y) = -w_B y + C\eta(\psi_{(C)}(y)) \int_0^\infty dJ(\zeta)\left[y\zeta/C + (\zeta/C)^2/2\right]$$

$$= -w_B y + \eta(\psi_{(C)}(y))\left[y + (1/2)m^{(2)}/C\right],$$

and then

$$\hat{G}_{(C)} g(y) \leq y(-w_B + \eta(\psi_{(C)}(y))) + m^{(2)}/(2C). \tag{6.5}$$

Note that, by our construction and (6.2), for all sufficiently large C (recall that $w_B \to v$), we have, for some fixed $\epsilon > 0$,

$$-w_B + \eta(\psi_C(y)) \geq \epsilon, \ y \leq -1, \quad \text{and} \quad -w_B + \eta(\psi_C(y)) \leq -\epsilon, \ y \geq 1. \tag{6.6}$$

From (6.4) we can see that, informally speaking, the operator $\hat{G}_{(C)}$ is an upper bound on the generator $G_{(C)}$ of the process $X_{(C)}(\cdot)$, when these operators are applied to the function g. This observation is only informal, because the function g is *not* even within the domain of $G_{(C)}$.

Formally, using (6.4) and the fact that the distribution $\psi_{(C)}$ is concentrated on $[-C_0, C_0]$, we can write

$$0 = \lim_{t \downarrow 0} \frac{1}{t} \psi_{(C)} \big[P_{(C)}^t - I \big] g \leq \lim_{t \downarrow 0} \frac{1}{t} \psi_{(C)} \big[\hat{P}_{(C)}^t - I \big] g = \psi_{(C)} \hat{G}_{(C)} g \tag{6.7}$$

and observe that

$$\psi_{(C)} \hat{G}_{(C)} g \leq -\epsilon \psi_{(C)} h_1 + C_2,$$

where $h_1(y) \doteq |y| I(|y| \geq 1)$, and C_2 does not depend on (sufficiently large) C. We obtain that, uniformly in large C,

$$\psi_{(C)} h_1 \leq C_2 / \epsilon.$$

This implies the tightness of $\{\psi_{(C)}\}$, thus completing the proof of the claim (6.3).

Given the tightness of $\{\psi_{(C)}\}$, consider a subsequence of C along which $\psi_{(C)} \xrightarrow{w} \psi_{(\infty)}$, where $\psi_{(\infty)}$ is a proper distribution, not necessarily continuous. Recall that $w_B \to v$. Then we can apply Lemma 6.4. We obtain that the entire distribution $\psi_{(\infty)}$ must be concentrated either (a) at a single point $z \in [-1, 1]$, or (b) on a segment $[z_1, z_2]$ of non-zero length, where $z_1 = \min\{y \mid \psi_{(\infty)}(y) = v_v\}$, $z_2 = \sup\{y \mid \psi_{(\infty)}(y) = v_v\}$, $v_v \doteq \eta^{-1}(v)$.

The case (a) is impossible, because the property that either $\psi_{(C)}^{-1}(v) = 1$ or $\psi_{(C)}^{-1}(1 - v) = -1$ cannot hold for all large C.

Let us show that the case (b) is also impossible. Suppose it holds. Then $z_1 \leq 0 \leq z_2$, because otherwise the median of $\psi_{(C)}$ could not stay at 0 for all C. We also know that (since $\psi_{(\infty)}(y) = v_v \in [z_1, z_2]$) the distribution $\psi_{(\infty)}$ consists of two atoms at z_1 and z_2, with positive weights v_v and $1 - v_v$, respectively. So, either $z_1 < 0$ or $z_2 > 0$. Consider, for concreteness, the case $z_2 > 0$. (The treatment of the case $z_1 < 0$ is analogous.) Fix small $\epsilon > 0$ and, for each C, consider the points

$$z_{1,\epsilon} = z_{1,\epsilon}(C) \doteq \psi_{(C)}^{-1}(v_v - \epsilon), \qquad z_{2,\epsilon} = z_{2,\epsilon}(C) \doteq \psi_{(C)}^{-1}(v_v + \epsilon).$$

Given the form of the limiting distribution $\psi_{(\infty)}$, as $C \to \infty$, we have

$$z_{1,\epsilon} \to z_1, \qquad z_{2,\epsilon} \to z_2.$$

Let $\epsilon_1 = \eta(v_v - \epsilon) - \eta(v_v + \epsilon)$. Obviously, $\epsilon_1 \downarrow 0$ as $\epsilon \downarrow 0$.

For any fixed ϵ (and corresponding ϵ_1), and arbitrarily small $\delta > 0$, the following holds for all sufficiently large C:

$$\hat{G}_{(C)} g(y) \leq y(-w_B + \eta(\psi_{(C)}(y))) + \delta \quad \text{(from (6.5))}. \tag{6.8}$$

We will now estimate $\psi_{(C)} \hat{G}_{(C)} g$ by considering intervals $(-\infty, z_{1,\epsilon}]$, $(z_{1,\epsilon}, z_{2,\epsilon}]$, $(z_{2,\epsilon}, \infty)$, and using the following: the $\psi_{(C)}$-measures of those intervals, estimates of values of $-w_B + \eta(\psi_{(C)}(y))$ over those intervals, and estimates of $|z_{1,\epsilon} - z_1|$ and $|z_{2,\epsilon} - z_2|$. The values and estimates listed just above are

$$|z_{1,\epsilon} - z_1| < \delta, \qquad |z_{2,\epsilon} - z_2| < \delta,$$

$$-w_B + \eta(\psi_{(C)}(y)) > \epsilon_1, \quad y \leq z_{1,\epsilon},$$

$$-w_B + \eta(\psi_{(C)}(y)) < \epsilon_1, \quad y \geq z_{2,\epsilon},$$

$$| - w_B + \eta(\psi_{(C)}(y))| \leq \epsilon_1, \quad z_{1,\epsilon} \leq y \leq z_{2,\epsilon},$$

$$\psi_{(C)}(z_{1,\epsilon}) = v_v - \epsilon, \qquad \psi_{(C)}(z_{2,\epsilon}) - \psi_{(C)}(z_{1,\epsilon}) = 2\epsilon, \qquad 1 - \psi_{(C)}(z_{2,\epsilon}) = 1 - v_v - \epsilon.$$

Then, from (6.8),

$$\psi_{(C)}\hat{G}_{(C)}g \le (\nu_v - \epsilon)(z_1 + \delta)\epsilon_1 + 2\epsilon[\max\{|z_1|, z_2\} + 2\delta]\epsilon_1 + (1 - \nu_v + \epsilon)(z_2 - \delta)(-\epsilon_1) + \delta$$
$$\le (\nu_v - \epsilon)\delta\epsilon_1 + 2\epsilon[\max\{|z_1|, z_2\} + 2\delta]\epsilon_1 + (1 - \nu_v + \epsilon)(z_2 - \delta)(-\epsilon_1) + \delta,$$

where the last inequality is because $z_1 \le 0$. If we pick sufficiently small $\epsilon > 0$ (with the corresponding, also small, $\epsilon_1 > 0$), and then choose sufficiently small δ, then the right-hand side in the last display is negative. This implies $\psi_{(C)}\hat{G}_{(C)}g < 0$, which is a contradiction with (6.7). Therefore, the case (b) is impossible.

Thus, the assumption that the family of distributions $\{\gamma_B\}$ is not tight leads to a contradiction. This completes the proof of the lemma. □

The following lemma completes the proof of Theorem 3.1(i).

Lemma 6.3. *Consider a subsequence of B. (We know that $w_B \to v$.) Then there exists a further subsequence along which γ_B converges to a proper distribution ϕ on $(-\infty, \infty)$, which is a traveling wave shape with the median at 0, $\phi(0) = 1/2$.*

Proof. Using the tightness of $\{\gamma_B\}$, we can find a further subsequence of B along which $\gamma_B \overset{w}{\to} \phi$; but, since the functions γ_B are uniformly Lipschitz in $(-B, \infty)$, we conclude that the function ϕ is $1/v$-Lipschitz, and then $\gamma_B \overset{u}{\to} \phi$. Also, clearly $\phi(0) = 1/2$.

If $P_B^t(y, H)$ is the transition function of the process $X_B(\cdot)$, then it is easy to see that, for each y and $t \ge 0$, $P_B^t(y, \cdot) \overset{w}{\to} P_\infty^t(y, \cdot)$, where $P_\infty^t(y, \cdot)$ is the transition function of the particle process $X_\infty(\cdot)$, corresponding to the distribution ϕ and speed v. Moreover, $P_\infty^t(y, \cdot)$ is continuous in y, i.e. the process $X_\infty(\cdot)$ is Feller continuous. Then (see, e.g., [9]) the limit ϕ of the stationary distributions γ_B of $X_B(\cdot)$ is the stationary distribution of the limiting process $X_\infty(\cdot)$. This means that ϕ is in fact the fixed point $\phi = A\phi$, and therefore, by Lemma 4, ϕ is a traveling wave shape. □

6.2. Proof of (6.4)

To prove (6.4) we first show that, for a small t, the contribution into the expectations $[P_{(C)}^t g](y)$ and $[\hat{P}_{(C)}^t g](y)$ of the events involving two or more particle jump urges in $[0, t]$ is $o(t)$. Indeed, consider for example $[\hat{P}_{(C)}^t g](y)$, which is equal to

$$\mathbb{E}[(\hat{X}_{(C)}(t))^2/2 | \hat{X}_{(C)}(0) = y].$$

We need to show that

$$\mathbb{E}[(\hat{X}_{(C)}(t))^2 | \hat{X}_{(C)}(0) = y; \text{ two or more jump urges in } [0, t]] \le o(t), \tag{6.9}$$

for a fixed C. Let U_i, $i = 1, 2, \ldots$, be an i.i.d. sequence of random variables with the distribution $J(\cdot)$ (of a jump size); let Π_t be a Poisson random variable, independent from the sequence, with mean t; let $\tilde{\Pi}_t = \Pi_t I\{\Pi_t \ge 2\}$. Note that $\mathbb{E}\tilde{\Pi}_t = t - te^{-t} = o(t)$, and $\mathbb{E}\tilde{\Pi}_t^2 = t + t^2 - te^{-t} = o(t)$. Define

$$Z = \left[\sum_{i=1}^{\Pi_t} U_i\right] I\{\Pi_t \ge 2\} = \sum_{i=1}^{\tilde{\Pi}_t} U_i,$$

and note that

$$\mathbb{E}Z \le \mathbb{E}U_1 \mathbb{E}\tilde{\Pi}_t = o(t), \qquad \mathbb{E}Z^2 = [\mathbb{E}\tilde{\Pi}_t^2 - \mathbb{E}\tilde{\Pi}_t](\mathbb{E}U_1)^2 + \mathbb{E}\tilde{\Pi}_t \mathbb{E}U_1^2 = o(t).$$

Finally, for bounded values of t, $t \in [0, a]$,

$$\mathbb{E}\big[(\hat{X}_{(C)}(t))^2 | \hat{X}_{(C)}(0) = y; \text{ two or more jump urges in } [0, t]\big]$$

$$\leq \mathbb{E}\left[C_0 + w_B a + \sum_{i=1}^{\Pi_t} U_i\right]^2 I\{\Pi_t \geq 2\}$$

$$\leq C_{11} \mathbb{E}I\{\Pi_t \geq 2\} + C_{12} \mathbb{E}Z + C_{13} \mathbb{E}Z^2 = o(t),$$

where the constants C_{11}, C_{12}, C_{13} do not depend on $y \in [-C_0, C_0]$. This proves (6.9).

We continue with the proof of (6.4). Consider the particle located at $y \in [-C_0, C_0]$ at time 0. In the interval $[0, t]$ we couple $X_{(C)}(\cdot)$ and $\hat{X}_{(C)}(\cdot)$ in the natural way so that, if the particle has at most one jump urge in this interval, then the particle jump times and corresponding jump sizes coincide. Then, if the particle (starting at y) has no jumps in $[0, t]$, $X_{(C)}^2(t) \leq \hat{X}_{(C)}^2(t)$ holds. Consider now the case when the particle has exactly one jump in $[0, t]$, at some time s. Note that $X_{(C)}(s) \neq \hat{X}_{(C)}(s)$, specifically $X_{(C)}(s) > \hat{X}_{(C)}(s)$, is only possible if the particle 'hits' the point $-C_0$ in $[0, s]$. Also note that, if the latter happens, $0 \leq \Delta \doteq X_{(C)}(s) - \hat{X}_{(C)}(s) \leq tw_B$, and then, using the relation $g(y + \delta) - g(y) = y\delta + \delta^2/2$, we can write

$$\mathbb{E}\big[g\big(X_{(C)}(s+)\big) - g\big(\hat{X}_{(C)}(s+)\big)\big] I\big\{X_{(C)}(s+) > \hat{X}_{(C)}(s+)\big\}$$

$$\leq \Delta \int_0^{2C_0} dJ(x)(x - C_0) + \Delta \bar{J}(2C_0) + o(t). \tag{6.10}$$

The integral in (6.10) is the upper bound on the expectation of the linear term of the difference $g\big(X_{(C)}(s+)\big) - g\big(\hat{X}_{(C)}(s+)\big)$ (the '$y\delta$') for the jump sizes $x \in [0, 2C_0]$; for such jump sizes, necessarily, $X_{(C)}(s+) > \hat{X}_{(C)}(s+)$. The $\Delta \bar{J}(2C_0)$ is the upper bound on the expected linear term of the difference $g\big(X_{(C)}(s+)\big) - g\big(\hat{X}_{(C)}(s+)\big)$ for jump sizes $x > 2C_0$; in this case $X_{(C)}(s+) = C_0 \geq \hat{X}_{(C)}(s+)$, and we can use the same bound as for jump sizes $x \leq C_0$. We can write

$$(6.10) \leq -\Delta \int_0^{C_0} dJ(x)(C_0 - x) + \Delta \int_{C_0}^{2C_0} dJ(x)(x - C_0) + \Delta \bar{J}(2C_0) + o(t)$$

$$= [-C_5 + C_6]\Delta + o(t),$$

where

$$C_5 = C_0 J(C_0) - \int_0^{C_0} dJ(x)x,$$

$$C_6 = \int_{C_0}^{2C_0} dJ(x)x - C_0[J(2C_0) - J(C_0)] + \bar{J}(2C_0).$$

Note that for all sufficiently large C_0, $C_5 > C_6$ holds, and then, for all sufficiently large C_0,

$$\mathbb{E}\big[g\big(X_{(C)}(s+)\big) - g\big(\hat{X}_{(C)}(s+)\big)\big] I\big\{X_{(C)}(s+) \geq \hat{X}_{(C)}(s+)\big\} \leq o(t). \tag{6.11}$$

Observe also that, if $X_{(C)}(s+) \geq \hat{X}_{(C)}(s+)$, which is right after the single jump, then

$$g(X_{(C)}(u)) - g(\hat{X}_{(C)}(u)) \text{ can only decrease in the interval } [s+, t], \tag{6.12}$$

because there are no jumps in it. From (6.11) and (6.12), for all sufficiently large C_0,

$$\mathbb{E}\big[g\big(X_{(C)}(t)\big) - g\big(\hat{X}_{(C)}(t)\big)\big] I\big\{X_{(C)}(s+) \geq \hat{X}_{(C)}(s+)\big\} \leq o(t). \tag{6.13}$$

Furthermore, if $\hat{X}_{(C)}(s+) > C_0$, and then necessarily $X_{(C)}(s+) = C_0$, we have

$$g\big(X_{(C)}(s+)\big) < g\big(\hat{X}_{(C)}(s+)\big)$$

and then also

$$g\big(X_{(C)}(t)\big) < g\big(\hat{X}_{(C)}(t)\big),$$

because, recall, we only consider small t. Thus,

$$\big[g\big(X_{(C)}(t)\big) - g\big(\hat{X}_{(C)}(t)\big)\big]I\big\{X_{(C)}(s+) < \hat{X}_{(C)}(s+)\big\} \leq 0 \quad \text{for small } t. \tag{6.14}$$

Combining (6.13) and (6.14), we finally obtain (6.4). $\qquad\square$

6.3. Supplementary fact: Lemma 6.4

The following supplementary fact, Lemma 6.4, is used in the proof of Lemma 6.2. It is used at the point where we assume that $C = C(B) \to \infty$ as $B \to \infty$, and the scaled distributions $\psi_{(C)}(y) = \gamma_B(Cy)$ are such that $\psi_{(C)} \xrightarrow{w} \psi_{(\infty)}$, where $\psi_{(\infty)}$ is a proper distribution; it is also known that the speed w_B converges, $w_B \to w_\infty$ (in fact, $w_\infty = v$). Lemma 6.4 basically says that as $C \to \infty$, the (fluid-scaled) process $X_{(C)}(t) = (1/C)X_B(Ct)$, whose stationary distribution is $\psi_{(C)}$, in the limit becomes (informally speaking) a deterministic (fluid-limit) process $X_{(\infty)}(t)$, whose stationary distribution is $\psi_{(\infty)}$. The process $X_{(\infty)}(t)$ is such that (informally speaking)

$$\frac{d}{dt}X_{(\infty)}(t) = \eta\big(\psi_{(\infty)}(X_{(\infty)}(t))\big) - w_\infty.$$

As a result, the distribution $\psi_{(\infty)}$ must be concentrated on those points y where (informally speaking) $\eta(\psi_{(\infty)}(y)) = w_\infty$.

Lemma 6.4. *Consider a subsequence of $B \to \infty$ and a function $C = C(B)$ such that $C \leq B$, $C \to \infty$, and $B/C \to C_0^* \leq \infty$. Assume that $w_B \to w_\infty$ for some constant $w_\infty \in (0, 1)$. Define $\psi_{(C)}(y) \doteq \gamma_B(Cy)$, and assume that $\psi_{(C)} \xrightarrow{w} \psi_{(\infty)}$, where $\psi_{(\infty)}$ is a proper distribution, not necessarily continuous. (It is concentrated on a finite interval $\big[-C_0^*, C_0^*\big]$ if $C_0^* < \infty$ or on the entirety of $\mathbb{R}$ otherwise.) For the chosen subsequence of B and the corresponding values C, consider the sequence of processes $X_{(C)}(t) \doteq (1/C)X_B(Ct)$, $t \geq 0$, with converging initial states $X_{(C)}(0) \to y_0 \in \mathbb{R}$. $\big($If $C_0^* < \infty$, then this necessarily means $y_0 \in \big[-C_0^*, C_0^*\big]$.$\big)$ Then the following hold:*

(i) *The sequence of processes can be constructed on a common probability space so that, w.p.1, any further subsequence of B has a still further subsequence of B along which $X_{(C)}(t) \to X_{(\infty)}(t)$, where $X_{(\infty)}(t)$ is a deterministic Lipschitz continuous trajectory, such that $X_{(\infty)}(0) = y_0$ and for all $t \geq 0$,*

$$\frac{d_\ell^+}{dt}X_{(\infty)}(t) \geq \eta\big(\psi_{(\infty)}\big(X_{(\infty)}(t)\big)\big) - w_\infty \quad \text{if } X_{(\infty)}(t) < C_0^*, \tag{6.15}$$

$$\frac{d_u^+}{dt}X_{(\infty)}(t) \leq \eta\big(\psi_{(\infty)}\big(X_{(\infty)}(t)-\big)\big) - w_\infty \quad \text{if } X_{(\infty)}(t) > -C_0^*. \tag{6.16}$$

(ii) *Any limiting trajectory $X_{(\infty)}(\cdot)$ in (i) is such that $X_{(\infty)}(t) \to \big[q^{low}, q^{up}\big]$, where*

$$q^{low} = \sup\{y \mid \eta(\psi_{(\infty)}(y)) > w_\infty\}, \qquad q^{up} = \inf\{y \mid \eta(\psi_{(\infty)}(y)) < w_\infty\}.$$

Moreover, this convergence is uniform in y_0 and all such $X_{(\infty)}(\cdot)$, as long as y_0 is restricted to a compact subset of $\mathbb{R}$.

(iii) *The distribution $\psi_{(\infty)}$ is concentrated on the segment $\left[q^{low}, q^{up}\right]$. Consequently, if $q^{low} < q^{up}$ (and then $\psi_{(\infty)}(y) = \eta^{-1}(w_\infty)$ for $y \in \left(q^{low}, q^{up}\right)$), the distribution $\psi_{(\infty)}$ has exactly two atoms, at the points q^{low} and q^{up}, with masses $\eta^{-1}(w_\infty)$ and $1 - \eta^{-1}(w_\infty)$, respectively.*

We remark that Lemma 6.4(i) easily implies a stronger property: w.p.1, $X_{(C)}(\cdot) \overset{u.o.c.}{\to} X_{(\infty)}(\cdot)$, where $X_{(\infty)}$ is the unique Lipschitz trajectory such that $X_{(\infty)}(0) = y_0$ and for almost all t with respect to Lebesgue measure,

$$X'_{(\infty)}(t) \in \left[\eta(\psi(X_{(\infty)}(t))) - w_\infty, \eta\left(\psi\left(X_{(\infty)}(t-)\right)\right) - w_\infty\right].$$

We do not need this stronger property in the present paper.

Proof of Lemma 6.4. The proof uses a fairly standard fluid-limit-type argument. When C is large, the process $X_{(C)}(\cdot)$ becomes 'almost deterministic'. Indeed, when C is large, and then the distribution $\psi_{(C)}$ is close to a proper distribution $\psi_{(\infty)}$, the process makes right jumps at the rate $O(C)$, with the jump sizes being $O(1/C)$; as a result, in the vicinity of time t the process trajectory is almost deterministic, with the derivative being close to $\eta(\psi_{(\infty)}\left(X_{(C)}(t)\right)) - w_\infty$. All properties stated in the lemma stem from this basic behavior.

We proceed with the formal proof. Consider the following natural common probability space construction for the particle processes X_B, for all B. A unit-rate Poisson process $\Pi(t)$, $t \geq 0$, drives the particle urges to jump. An i.i.d. sequence $Z_1, Z_2, \ldots$ of random variables with distribution $J(\cdot)$ determines the sequence of particle jump sizes, when it does jump. An i.i.d. sequence $\Xi_1, \Xi_2, \ldots$ of random variables uniformly distributed in $[0, 1]$ determines whether or not the particle actually jumps when it gets the ith urge to jump; specifically, if a particle gets the ith urge to jump at time t, when its location is $y = X_B(t)$, it actually jumps if $\Xi_i \leq \bar{\eta}(y, \gamma_B)$. This construction ensures that, for each B, the process X_B is (up to stochastic equivalence) as defined.

The driving sequences satisfy the functional strong law of large numbers (FSLLN) properties: as $B \to \infty$ (and the corresponding $C = C(B) \to \infty$), w.p.1,

$$\left(\frac{1}{C}\Pi(Ct), t \geq 0\right) \overset{u.o.c.}{\to} (t, t \geq 0), \tag{6.17}$$

$$\left(\frac{1}{C}\sum_{i=1}^{\lfloor Cs \rfloor} Z_i, s \geq 0\right) \overset{u.o.c.}{\to} (s, s \geq 0), \tag{6.18}$$

$$\left(\frac{1}{C}\sum_{i=1}^{\lfloor Cs \rfloor} I\{\Xi_i \leq \xi\}, 0 \leq \xi \leq 1, s \geq 0\right) \overset{u.o.c.}{\to} (\xi s, 0 \leq \xi \leq 1, s \geq 0). \tag{6.19}$$

(i) We have $X_{(C)}(t) = X_{(C)}(0) + X_{(C)}^{\uparrow}(t) - X_{(C)}^{\downarrow}(t)$, where $X_{(C)}^{\uparrow}(t)$ accounts for the forward (right) jumps in $[0, t]$ and $X_{(C)}^{\downarrow}(t)$ accounts for the total distance traveled by the particle backwards (left) in $[0, t]$ (at a speed which may be either w_B or 0). Both $X_{(C)}^{\uparrow}$ and $X_{(C)}^{\downarrow}$ are non-decreasing; $X_{(C)}^{\downarrow}$ is w_B-Lipschitz; given the FSLLN properties (6.17)–(6.19), we observe that the sequence of processes $X_{(C)}^{\uparrow}$ is asymptotically Lipschitz, meaning, w.p.1, for any $0 \leq t_1 \leq t_2 < \infty$,

$$\limsup_{C \to \infty} \left[X_{(C)}^{\uparrow}(t_2) - X_{(C)}^{\uparrow}(t_1)\right] \leq t_2 - t_1.$$

This implies that, w.p.1., any subsequence of trajectories $X_{(C)}(\cdot)$ has a further subsequence along which

$$X_{(C)}(\cdot) \overset{u.o.c.}{\to} X_{(\infty)}(\cdot),$$

where $X_{(\infty)}(t)$ is $(1 + w_\infty)$-Lipschitz, and $X_{(\infty)}(0) = y_0$. It remains to show that a limit trajectory $X_{(\infty)}(\cdot)$ satisfies (6.15)–(6.16). This easily follows from (6.17)–(6.19). Consider (6.15), for example. If $X_{(\infty)}(t) = y$, then for an arbitrarily small $\epsilon > 0$ there exists $\delta > 0$ such that for all $t' \in (t - \delta, t + \delta)$ and all sufficiently large C we have

$$\psi_{(C)}\big(X_{(C)}(t')\big) < \psi_{(\infty)}(y) + \epsilon$$

and then

$$\bar{\eta}\big(X_{(C)}(t'), \psi_{(C)}\big) > \eta\big(\psi_{(\infty)}(y) + \epsilon\big).$$

In other words, when t' is close to t and C is sufficiently large, the probability of the particle of $X_{(C)}$ jumping when it gets an urge is bounded below by $\eta(\psi_{(\infty)}(X_{(\infty)}(t)) + \epsilon)$. Application of (6.17)–(6.19) easily gives (6.15). We omit further details. The property (6.16) is shown similarly.

(ii) This follows from (i).

(iii) This follows easily from (ii). Indeed, (ii) implies that for any compact set $K \in \mathbb{R}$ there exists $T > 0$ such that, for all sufficiently large C, with uniformly high probability, if $X_{(C)}(0) \in K$ then $X_{(C)}(T)$ is close to $\big[q^{low}, q^{up}\big]$. Therefore, the stationary distribution of $X_{(C)}(\cdot)$, i.e. $\psi_{(\infty)}$, must in the limit concentrate on $\big[q^{low}, q^{up}\big]$. $\qquad\square$

6.4. Proof of Theorem 3.1(ii): finite moments

The distribution ϕ is proper and $\phi = A\phi$. Then (3.3) follows from Lemma 4.1, (4.1). $\qquad\square$

6.5. Proof of Theorem 3.2

If the assumption (3.4) holds, the mean-field model $f(x, t)$ is within the assumptions of that in [7]. (Actually, in [7] it is assumed that $J(y) = 1 - e^{-y}$ and $\eta(\nu) = (1 - \nu)^K$, $K \geq 1$. However, all proofs in [7] hold *as is* under the more general assumption (3.4) on $J(\cdot)$, and for arbitrary continuous strictly decreasing $\eta(\cdot)$.) Theorem 1 of [7] states that *if a traveling wave shape ϕ with well-defined finite mean exists*, then the statements (i) and (ii) of our Theorem 3.2 hold. (Note that the uniform convergence in (3.5) follows from the L_1-convergence, because $\phi(\cdot)$ is continuous and non-decreasing, and each $f(\cdot + vt, t)$ is non-decreasing.) But, by our Theorem 3.1, such a traveling wave shape does exist, which completes the proof. $\qquad\square$

7. Discussion

7.1. A generalization: more general jump process

Our main results (along with proofs) easily extend, for example, to the following system. There are n particles, moving in jumps in the positive direction. Each particle can make two types of jumps. Jumps of type 1 are those in our original model: they are driven by an independent Poisson process of rate $\mu \geq 0$, the jump probabilities are given by a non-increasing function $\eta_n(\cdot)$ uniformly converging to a strictly decreasing continuous function $\eta(\nu)$, and the jump sizes are independent with distribution $J(\cdot)$. Suppose that, in addition, each particle can make jumps of type 2: they are driven by an independent Poisson process of rate $\mu_2 \geq 0$,

at which points the particle jumps right w.p.1, and the jump sizes are independent with distribution $J_2(y)$, $y \geq 0$; we define $\bar{J}_2(y) = 1 - J_2(y)$. Assume that (1.1) holds for $J(\cdot)$, and the analogous assumption holds for $J_2(\cdot)$ as well, for the same integer $\ell \geq 2$ as in (1.1):

$$m_2^{(k)} \doteq \int_0^\infty y^k dJ_2(y) < \infty, \qquad k = 1, 2, \ldots, \ell. \tag{7.1}$$

So, $m_2^{(1)} < \infty$ is the mean type-2 jump size.

This generalization may be useful for modeling situations when each particle may change its state either independently (type 2 jumps) or depending on the locations of other particles (type 1 jumps).

The corresponding mean-field model is still given by Definition 3.1, except that (3.1) generalizes to

$$\frac{\partial}{\partial t} f(x, t) = -\mu \int_{-\infty}^x d_y f(y, t) \bar{\eta}(y, f(\cdot, t)) \bar{J}(x - y) - \mu_2 \int_{-\infty}^x d_y f(y, t) \bar{J}_2(x - y). \tag{7.2}$$

The speed v generalizes to

$$v \doteq \mu m^{(1)} \int_0^1 \eta(\xi) d\xi + \mu_2 m_2^{(1)},$$

for which the 'conservation law' (3.2) holds. Therefore, just as in the original model, if the initial mean $\bar{f}(0)$ is well-defined and finite, then it moves right at the constant speed v: $\bar{f}(t) = \bar{f}(0) + vt$.

A traveling wave shape is given by Definition 3.2 and is easily shown to satisfy a more general form of (1.3),

$$v\phi'(x) = \mu \int_{-\infty}^x \phi'(y) \eta(\phi(y)) \bar{J}(x - y) dy + \mu_2 \int_{-\infty}^x \phi'(y) \bar{J}_2(x - y) dy, \tag{7.3}$$

for each x; the derivative $\phi'(x)$ is continuous.

For this more general mean-field model, *Theorem 3.1 (with the condition (1.1) complemented by (7.1)) and Theorem 3.2 (with the condition (3.4) complemented by the analogous condition on $J_2(\cdot)$) hold as they are.* The proofs are same, up to straightforward adjustments.

The key feature that the more general model shares with the original one—and which makes the same analysis work—is that jumps of a particle depend only on its location quantile. (This, in particular, implies that the speed of the mean $\bar{f}(t)$, and then of a traveling wave $\phi(x - vt)$, is known in advance and equal to v.) As long as this key feature is preserved for other mean-field models, we believe that our main results and analysis have a good chance of extending to such other models as well.

7.2. The condition (3.4) on the jump size distribution

The condition (3.4) is used in [7] to prove that the L_1-distance between any two mean-field models (with equal mean) is strictly decreasing as long as these mean-field models are different. (Actually, the results in [7] are specifically for the exponential $J(\cdot)$, but as far as the decreasing L_1-distance is concerned, only the condition (3.4) on $J(\cdot)$ is used.) The decreasing L_1-distance result of [7] is then used in the proof of our Theorem 3.2.

It is likely that the analysis in [7] can be generalized to establish the decreasing L_1-distance property under a *much relaxed condition* (3.4). If so, our Theorem 3.2 holds under a relaxed condition (3.4) as well.

7.3. A conjecture about the limit of stationary distribution

Consider the stochastic particle system with n particles (not the corresponding mean-field model, which is the focus of this paper). Recall that by $f^n(\cdot, t)$ we denote its random state (the empirical distribution of particle locations) at time t. Denote by $f_*^n(\cdot, t)$ the function $f^n(\cdot, t)$, recentered so that its median is at 0. Assume (1.1) and (3.4). It is not hard to see (using the fluid limit technique, for example) that, for any fixed n, the process $f_*^n(\cdot, t)$ is stochastically stable (positive Harris recurrent), and therefore has unique stationary distribution. (We do not give the details of the stability proof—which are not hard—because analysis of the stochastic particle system with finite n is not the subject of this paper.) By our Theorems 3.1 and 3.2, there exists a unique (up to a shift) traveling wave shape $\phi(\cdot)$. Then, the results of [6] (convergence of $f^n(\cdot, t)$ to a mean-field model $f(\cdot, t)$, as $n \to \infty$) and our Theorem 3.2 (convergence of a mean-field model $f(\cdot, t)$ to the traveling wave, as $t \to \infty$) strongly suggest the following conjecture about the limit of stationary distributions of $f_*^n(\cdot, t)$.

Conjecture 7.1. *Assume* (1.1) *and* (3.4). *Let $f_*^n(\cdot, \infty)$ be a random value of $f_*^n(\cdot, t)$ in the stationary regime. Let $\phi(\cdot)$ be the unique traveling wave shape, centered to have the median at 0. Then, as $n \to \infty$, $f_*^n(\cdot, \infty)$ concentrates at $\phi(\cdot)$; that is,*

$$\|f_*^n(\cdot, \infty) - \phi(\cdot)\| \xrightarrow{P} 0.$$

As natural as it is, this conjecture (which is also supported by the simulation experiments in [7]) does *not* directly follow from the results of [6] and our Theorem 3.2. Establishing Conjecture 7.1 is a subject for future work.

Funding information

Competing interests

There were no competing interests to declare which arose during the preparation or publication process for this article.

References

[1] BALÁZS, M., RÁCZ, M. Z. AND TÓTH, B. (2014). Modeling flocks and prices: jumping particles with an attractive interaction. *Ann. Inst. H. Poincaré Prob. Statist.* **50**, 425–454.

[2] BRAMSON, M. (2008). *Stability of Queueing Networks*. Springer, Berlin, Heidelberg.

[3] DAI, J. G. (1995). On the positive Harris recurrence for multiclass queueing networks: a unified approach via fluid models. *Ann. Appl. Prob.* **5**, 49–77.

[4] DAI, J. G. AND MEYN, S. P. (1995). Stability and convergence of moments for multiclass queueing networks via fluid limit models. *IEEE Trans. Automatic Control* **40**, 1889–1904.

[5] ETHIER, S. AND KURTZ, T. (1986). *Markov Processes: Characterization and Convergence*. John Wiley, New York.

[6] GREENBERG, A., MALYSHEV, V. AND POPOV, S. (1997). Stochastic model of massively parallel computation. *Markov Process. Relat. Fields* **2**, 473–490.

[7] GREENBERG, A., SHENKER, S. AND STOLYAR, A. (1996). Asynchronous updates in large parallel systems. *ACM SIGMETRICS Performance Evaluation Rev.* **24**, 91–103.

[8] KANTOROVICH, L. V. AND AKILOV, G. P. (1982). *Functional Analysis*, 2nd edn. Pergamon Press, Oxford.

[9] LIPTSER, R. S. AND SHIRYAEV, A. N. (1989). *Theory of Martingales*. Kluwer Academic Publishers, Dordrecht.

[10] MALYSHEV, V. AND MANITA, A. (2006). Phase transitions in the time synchronization model. *Theory Prob. Appl.* **50**, 134–141.

[11] MALYSHKIN, A. (2006). Limit dynamics for stochastic models of data exchange in parallel computation networks. *Problems Inf. Transmission* **42**, 234–250.

[12] MANITA, A. (2006). Markov processes in the continuous model of stochastic synchronization. *Russian Math. Surveys* **61**, 993–995.

[13] MANITA, A. (2009). Stochastic synchronization in a large system of identical particles. *Theory Prob. Appl.* **53**, 155–161.

[14] MANITA, A. (2014). Clock synchronization in symmetric stochastic networks. *Queueing Systems* **76**, 149–180.

[15] MANITA, A. AND SHCHERBAKOV, V. (2005). Asymptotic analysis of a particle system with mean-field interaction. *Markov Process. Relat. Fields* **11**, 489–518.

[16] RYBKO, A. AND STOLYAR, A. (1992). Ergodicity of stochastic processes describing the operation of open queueing networks. *Problems Inf. Transmission* **28**, 199–200.

[17] STOLYAR, A. L. (1995). On the stability of multiclass queueing networks: a relaxed sufficient condition via limiting fluid processes. *Markov Process. Relat. Fields* **1**, 491–512.