Electronic Journal of Probability

Electron. J. Probab. **29** (2024), article no. 199, 1–24.
ISSN: 1083-6489 https://doi.org/10.1214/24-EJP1253

A two-line representation of stationary measure for open TASEP*

Włodzimierz Bryc† Pavel Zatitskii†

Abstract

We show that the stationary measure for the totally asymmetric simple exclusion process on a segment with open boundaries is given by a marginal of a two-line measure with a simple and explicit description. We use this representation to analyze asymptotic fluctuations of the height function near the triple point for a larger set of parameters than was previously studied. As a second application, we determine a single expression for the rate function in the large deviation principle for the height function in the fan and in the shock region. We then discuss how this expression relates to the expressions for the rate function available in the literature.

Keywords: Gibbs line measure; totally asymmetric exclusion process; large deviations; fluctuations of particle density.
MSC2020 subject classifications: 60K35; 60F10.
Submitted to EJP on July 19, 2024, final version accepted on December 5, 2024.

1 Introduction

A totally asymmetric simple exclusion process (TASEP) with open boundaries is a continuous-time finite-state Markov process that models the movement of particles along the N sites $\{1,\ldots,N\}$ from the left reservoir to the right reservoir. The particles cannot occupy the same site, and can move only to the nearest site on the right at rate 1. The particles arrive at the first location, if empty, at rate $\alpha>0$ and leave the system from the N-th site, if occupied, at rate $\beta>0$. For a description of the infinitesimal generator of this Markov process, we refer to e.g. [23, Section 3]. We will be interested solely in the stationary measure of this process.

The stationary measure for open TASEP has been studied for a long time, with explicit expressions available in [25] and [20]. In this paper we establish a two-line representation for this stationary measure in terms of a pair of weighted random walks. We remark that there are numerous other representations for the stationary measure

*This work was supported by a Simons grant (703475, WB).
†Department of Mathematical Sciences, University of Cincinnati, 2815 Commons Way, Cincinnati, OH, 45221-0025, USA. E-mail: wlodek.bryc@gmail.com,zatitspl@ucmail.uc.edu

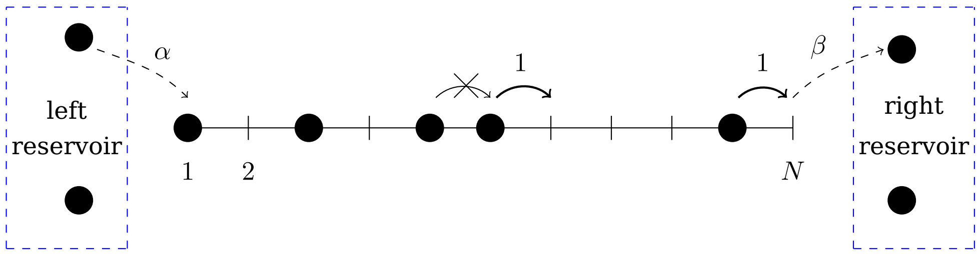

Figure 1: Totaly Asymmetric Simple Exclusion process with boundary parameters α, β.

of TASEP; a representation in [24, Section 5.2] does not separate the "two lines", but it covers a more general ASEP. Integral representation of the probability generating function [8, 26] is useful in studying Laplace transformations of limiting fluctuations. Ref. [21, Section 3.2] represents the stationary measure of a sequential TASEP as a marginal of a "two-layer" configuration that has a different form than ours.

The Gibbs measure (or line ensemble) representations have been valuable in studying integrable probabilistic models on full or half-space and have been extended to time-homogeneous models on an interval with two-sided boundary conditions in [1]. Barraquand, Corwin, and Yang [1, Theorem 1.3] establish that stationary measures for the free-energy increment process of geometric last passage percolation on a diagonal strip are described as marginals of explicitly defined two-layer Gibbs measures. They further pose the question of obtaining an explicit description of the open TASEP stationary measure from its implicit connection to the stationary solution of the exponential large passage percolation recurrence. This paper proposes an alternative approach implicitly based on the matrix method [14]. We demonstrate that a representation akin to their two-layer Gibbs measure holds for the stationary measure of the TASEP. We then use this representation to analyze asymptotic fluctuations of particle density for a larger set of parameters than was previously studied, and to prove the large deviations principle with a single expression for the rate function valid for all $\alpha, \beta \in (0, 1)$.

We now introduce configuration spaces and probability measures that will facilitate canonical representations of the random variables that we need. We begin with the stationary measure of TASEP which defines a (discrete) probability measure $\mathbb{P}_{\mathrm{TASEP}}$ on the configuration space $\Omega = \Omega^{(N)} = \{0, 1\}^N$. (We will omit the superscript N when it is fixed in an argument and clearly recognizable from the context.) We assign probability $\mathbb{P}_{\mathrm{TASEP}}(\boldsymbol{\tau})$ to a sequence $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_N) \in \Omega$ that encodes the occupied and empty sites, where $\tau_j \in \{0, 1\}$ is the occupation indicator of the j-th site. It will be convenient to parameterize $\mathbb{P}_{\mathrm{TASEP}}$ using

$$\mathsf{a} = \frac{1-\alpha}{\alpha}, \quad \mathsf{b} = \frac{1-\beta}{\beta}. \tag{1.1}$$

Formula (1.1) makes sense for all $\alpha, \beta > 0$, and then $\mathbb{P}_{TASEP}$ is determined by $\mathsf{a}, \mathsf{b} > -1$ in formula (2.1), but in this paper we will only consider $\alpha, \beta \in (0, 1)$, so that $\mathsf{a}, \mathsf{b} > 0$.

The steady state *height function* $\boldsymbol{H}_N$ is defined by

$$H_N(k) = \tau_1 + \cdots + \tau_k, \quad k = 0, 1, \ldots, N. \tag{1.2}$$

The invariant law $\mathbb{P}_{\mathrm{TASEP}}$ is uniquely determined by the law $\mathbb{P}_{\mathrm{H}}^{(N)}$ induced by $\boldsymbol{H}_N$ on the configuration space

$$\mathcal{S} := \{\boldsymbol{s} = (s_j)_{0 \le j \le N} : \ s_0 = 0, \ s_j - s_{j-1} \in \{0, 1\}, \ 1 \le j \le N\}.$$

Indeed,

$$\mathbb{P}_{\mathrm{H}}^{(N)}(\boldsymbol{s}) = \mathbb{P}_{\mathrm{TASEP}}(\boldsymbol{\tau})$$

with unique $\boldsymbol{\tau}$ such that $\boldsymbol{s} = \boldsymbol{H}_N(\boldsymbol{\tau})$, as $\boldsymbol{H}_N : \Omega \to \mathcal{S}$ is a bijection. Instead of determining $\mathbb{P}_{\rm TASEP}$, we will therefore determine $\mathbb{P}_{\rm H}^{(N)}$ as a marginal law of the top line of the two-line ensemble on the configuration space $\mathcal{S} \times \mathcal{S}$.

Denote by $\mathbb{P}_{\rm rw}$, the uniform law on $\mathcal{S} \times \mathcal{S}$ defined by two independent random walks with i. i. d. Bernoulli increments $\Pr(\xi = 0) = \Pr(\xi = 1) = 1/2$,

$$\mathbb{P}_{\rm rw}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \Pr(\boldsymbol{S}_1 = \boldsymbol{s}_1) \Pr(\boldsymbol{S}_2 = \boldsymbol{s}_2) = 1/4^N, \quad \boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathcal{S}.$$

The *two-line ensemble* (TLE) is the probability measure $\mathbb{P}_{\rm TLE}$ on $\mathcal{S} \times \mathcal{S}$, defined as follows:

$$\mathbb{P}_{\rm TLE}\left(\boldsymbol{s}_1, \boldsymbol{s}_2\right) = \frac{1}{\mathfrak{C}_{\mathsf{a},\mathsf{b}}} \frac{\mathsf{b}^{s_1(N)-s_2(N)}}{(\mathsf{ab})^{\min_{0 \le j \le N}\{s_1(j)-s_2(j)\}}} \mathbb{P}_{\rm rw}(\boldsymbol{s}_1, \boldsymbol{s}_2), \tag{1.3}$$

where $\mathfrak{C}_{\mathsf{a},\mathsf{b}}$ is the normalization constant and $\boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathcal{S}^{(N)}$. We will write $\mathbb{E}_{\rm TLE}$ for the expected value with respect to $\mathbb{P}_{\rm TLE}$. Of course, $\mathbb{P}_{\rm TLE} = \mathbb{P}_{\rm TLE}^{(N)}$ depends on N.

The two canonical coordinate mappings $\boldsymbol{S}_1, \boldsymbol{S}_2 : \mathcal{S} \times \mathcal{S} \to \mathcal{S}$ given by

$$\boldsymbol{S}_1(\boldsymbol{s}_1, \boldsymbol{s}_2) = \boldsymbol{s}_1, \quad \boldsymbol{S}_2(\boldsymbol{s}_1, \boldsymbol{s}_2) = \boldsymbol{s}_2, \tag{1.4}$$

give a canonical realization on $(\mathcal{S} \times \mathcal{S}, \mathbb{P}_{\rm rw})$ of the pair of independent Bernoulli random walks $\boldsymbol{S}_1, \boldsymbol{S}_2$ and at the same time they give a realization of the two-line ensemble on $(\mathcal{S} \times \mathcal{S}, \mathbb{P}_{\rm TLE})$.

Our main result is the following TASEP analog of [1, Theorem 1.3].

Theorem 1.1. *For* $\mathsf{a}, \mathsf{b} > 0$*, the marginal law of random sequence* $\boldsymbol{S}_1$ *given by* (1.4) *under measure* $\mathbb{P}_{\rm TLE}$ *is the (unique) law of the steady state height function* $\boldsymbol{H}_N$ *of the TASEP with parameters* α, β *given by* (1.1)*. That is, for* $\boldsymbol{s} \in \mathcal{S}$*,*

$$\mathbb{P}_{\rm H}(\boldsymbol{s}) = \sum_{\boldsymbol{s}' \in \mathcal{S}} \mathbb{P}_{\rm TLE}(\boldsymbol{s}, \boldsymbol{s}'). \tag{1.5}$$

Remark 1.2. As pointed out to us by an anonymous reviewer, Theorem 1.1 in fact holds for $\mathsf{a}, \mathsf{b} \ge 0$, i.e., for $\alpha, \beta \in (0, 1]$. The only change that is required is to rewrite (1.3) as an expression in a, b with non-negative exponents:

$$\frac{1}{\mathfrak{C}_{\mathsf{a},\mathsf{b}}} \mathsf{b}^{s_1(N)-s_2(N)-\min_{0 \le j \le N}\{s_1(j)-s_2(j)\}} \mathsf{a}^{\max_{0 \le j \le N}\{s_2(j)-s_1(j)\}} \mathbb{P}_{\rm rw}(\boldsymbol{s}_1, \boldsymbol{s}_2).$$

Then the identities that we establish in the proof for $\mathsf{a}, \mathsf{b} > 0$ extend to $\mathsf{a}, \mathsf{b} \ge 0$ by continuity.

The proof of Theorem 1.1 appears in Section 2. In Section 3 we give two applications which show how Theorem 1.1 allows to deduce asymptotic of the height function of TASEP from well known asymptotic properties of random walks. In Theorem 3.1 we use Theorem 1.1 and Donsker's theorem to obtain convergence of the fluctuations of TASEP to the process conjectured to be a stationary measure of a KPZ fixed point on an interval in the full range of parameters. To our knowledge, previously available results of this form, see [7, Theorem 1.5], required that the sum $\mathsf{u} + \mathsf{v}$ of the parameters in (3.2) be non-negative. (On the other hand, the more general five-parameter ASEP was covered.) In Theorem 3.4 we show that in the case of TASEP the large deviation principle for the height function is a consequence of Theorem 1.1, Mogulskii's theorem, and the contraction principle. The large deviation principle for the height function of a more general ASEP has been analyzed in [16], but besides the simplicity of the proof, a slight novelty here is the unified proof and an expression for the rate function, which works for all $\alpha, \beta \in (0, 1)$, i.e., for all $\mathsf{a}, \mathsf{b} > 0$, in the so called fan region $\mathsf{ab} < 1$ and in the shock region $\mathsf{ab} > 1$. Since the relation of our rate function to formulas [16, (1.7), (3.3), and (1.11)] is not obvious, we discuss this topic in Section 4.

2 Proof of Theorem 1.1

For $\mathsf{ab} < 1$, Theorem 1.1 can be deduced from [3, Section 2.3] used with $q = 0$. The general case can then be obtained by analytic continuation as discussed in [1, Remark 1.9]. Our more direct proof is based on induction on the size $N = 1, 2, \dots$ of the system and relies on a recursion for the invariant probabilities. Recursions for the invariant probabilities of a more general open asymmetric simple exclusion process (ASEP) appear in [23], [11], and [13]. Here, we will use a recursion that arises directly from the celebrated matrix method developed in [20]. This recursion appears under the name *basic weight equations* in [6, Theorem 1] and it has already been used for similar purposes in [24, Section 2.2]. Specified to TASEP, the basic weight equations say that the unique stationary measure of TASEP under reparameterization (1.1) is given by

$$\mathbb{P}_{\mathrm{TASEP}}(\boldsymbol{\tau}) = \frac{1}{\mathsf{Z}_{\mathsf{a},\mathsf{b}}} p_N(\boldsymbol{\tau}), \quad \boldsymbol{\tau} \in \Omega^{(N)}, \tag{2.1}$$

where $\{p_N(\boldsymbol{\tau})\}$ satisfy the recursion that determines the un-normalized steady state weights p_N uniquely in terms of the steady state weights p_{N-1} for a TASEP on $\Omega^{(N-1)}$. With the initial conditions

$$p_1(0) = 1 + \mathsf{a}, \quad p_1(1) = 1 + \mathsf{b}, \tag{2.2}$$

for $N \geq 2$ the recursion is:

$$p_N(0, \tau_2, \dots, \tau_N) = (1 + \mathsf{a}) p_{N-1}(\tau_2, \dots, \tau_N), \tag{2.3}$$

$$p_N(\tau_1, \dots, \tau_{N-1}, 1) = (1 + \mathsf{b}) p_{N-1}(\tau_1, \dots, \tau_{N-1}). \tag{2.4}$$

$$\begin{aligned} p_N(\tau_1, \dots, \tau_{n-1}, 1, 0, \tau_{n+2}, \dots, \tau_N) &= p_{N-1}(\tau_1, \dots, \tau_{n-1}, 1, \tau_{n+2}, \dots, \tau_N) \\ &\quad + p_{N-1}(\tau_1, \dots, \tau_{n-1}, 0, \tau_{n+2}, \dots, \tau_N), \end{aligned} \tag{2.5}$$

$1 \leq n \leq N - 1$.

2.1 The key identity and the proof of Theorem 1.1

We introduce a family of functions $f_N : \Omega^{(N)} \to (0, \infty)$ that we will use to prove (1.5). For $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N) \in \Omega^{(N)}$ and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_N) \in \Omega^{(N)}$, let

$$s_1(k) = \sum_{j=1}^k \tau_j, \quad s_2(k) = \sum_{j=1}^k \xi_j \tag{2.6}$$

with $s_1(0) := 0$ and $s_2(0) := 0$. Formula (2.6) defines a pair of bijections $\Omega^{(N)} \to \mathcal{S}$ and throughout this proof we will treat $\boldsymbol{s}_1 = \boldsymbol{s}_1(\boldsymbol{\tau})$ as a function of $\boldsymbol{\tau}$ and $\boldsymbol{s}_2 = \boldsymbol{s}_2(\boldsymbol{\xi})$ as a function of $\boldsymbol{\xi}$.

With the above convention, we introduce

$$f_N(\boldsymbol{\tau}) = f_N(\tau_1, \dots, \tau_N) = \sum_{(\xi_1, \dots, \xi_N) \in \Omega^{(N)}} \frac{\mathsf{b}^{s_1(N) - s_2(N)}}{(\mathsf{ab})^{\min_{0 \leq j \leq N} \{s_1(j) - s_2(j)\}}}. \tag{2.7}$$

Theorem 1.1 is a consequence of the following identity:

Lemma 2.1. *For $(\tau_1, \dots, \tau_N) \in \Omega^{(N)}$ and $N \geq 1$, we have*

$$f_N(\tau_1, \dots, \tau_N) = p_N(\tau_1, \dots, \tau_N), \tag{2.8}$$

with $p_N(\boldsymbol{\tau})$ from representation (2.1)*.*

Proof. It is clear that (2.8) holds for $N=1$. Indeed, in this case (2.7) is the sum of two terms corresponding to $\xi_1=0$ and $\xi_1=1$:

$$f_1(0)=\frac{\mathsf{b}^0}{(\mathsf{ab})^0}+\frac{\mathsf{b}^{-1}}{(\mathsf{ab})^{-1}}=1+\mathsf{a},$$
$$f_1(1)=\frac{\mathsf{b}^1}{(\mathsf{ab})^0}+\frac{\mathsf{b}^0}{(\mathsf{ab})^0}=\mathsf{b}+1,$$

matching the initial conditions (2.2).

Next, we show that f_N satisfies the same three recursions as p_N for $N\geq 2$. Throughout this proof, for $k=0,\dots,N$ we consider partial sums $\boldsymbol{s}_1\in\mathcal{S}^{(N)}$ and $\widetilde{\boldsymbol{s}}_1\in\mathcal{S}^{(N-1)}$ that depend on $\tau_1,\dots,\tau_n$ and partial sums $\boldsymbol{s}_2\in\mathcal{S}^{(N)}$ and $\widetilde{\boldsymbol{s}_2}\in\mathcal{S}^{(N-1)}$ that depend on the auxiliary $\{0,1\}$-valued variables $\xi_1,\dots,\xi_n$ that appear under the sum in (2.7). In one place in the last part of the proof, the sequence $\widetilde{\boldsymbol{s}}_1$ will be an explicit function of both sequences $\boldsymbol{\tau}$ and $\boldsymbol{\xi}$. We express f_N in terms of $\boldsymbol{s}_1,\boldsymbol{s}_2\in\mathcal{S}^{(N)}$ and relate it to f_{N-1} written in terms of $\widetilde{\boldsymbol{s}}_1,\widetilde{\boldsymbol{s}}_2\in\mathcal{S}^{(N-1)}$.

First, we verify that

$$f_N(0,\tau_2,\dots,\tau_N)=(1+\mathsf{a})f_{N-1}(\tau_2,\dots,\tau_N).$$

To see this, we define $\widetilde{s}_1(k)=\sum_{j=1}^k\tau_{j+1}$ and $\widetilde{s}_2(k)=\sum_{j=1}^k\xi_{j+1}$ so that $s_1(N)=\widetilde{s}_1(N-1)$, $s_2(N)=\xi_1+\widetilde{s}_2(N-1)$ and $\min_{0\leq j\leq N}\{s_1(j)-s_2(j)\}=-\xi_1+\min_{0\leq j\leq N-1}\{\widetilde{s}_1(j)-\widetilde{s}_2(j)\}$. Then (2.7) gives

$$\begin{aligned}f_N(0,\tau_2,\dots,\tau_N)&=\sum_{(\xi_1,\xi_2,\dots,\xi_N)\in\{0,1\}^N}\frac{\mathsf{b}^{-\xi_1}}{(\mathsf{ab})^{-\xi_1}}\cdot\frac{\mathsf{b}^{\widetilde{s}_1(N-1)-\widetilde{s}_2(N-1)}}{(\mathsf{ab})^{\min_{0\leq j\leq N-1}\{\widetilde{s}_1(j)-\widetilde{s}_2(j)\}}}\\&=\sum_{\xi_1\in\{0,1\}}\mathsf{a}^{\xi_1}\sum_{(\xi_2,\dots,\xi_N)\in\{0,1\}^{N-1}}\frac{\mathsf{b}^{\widetilde{s}_1(N)-\widetilde{s}_2(N)}}{(\mathsf{ab})^{\min_{0\leq j\leq N-1}\{\widetilde{s}_1(j)-\widetilde{s}_2(j)\}}}\\&=(1+\mathsf{a})f_{N-1}(\tau_2,\dots,\tau_N).\end{aligned}$$

Next, by a similar argument we verify that

$$f_N(\tau_1,\dots,\tau_{N-1},1)=(1+\mathsf{b})f_{N-1}(\tau_1,\dots,\tau_{N-1}).$$

In this case, we introduce $\widetilde{s}_1(k)=\sum_{j=1}^k\tau_j$ and $\widetilde{s}_2(k)=\sum_{j=1}^k\xi_j$ so that $s_1(N)=\widetilde{s}_1(N-1)+1$, $s_2(N)=\widetilde{s}_2(N-1)+\xi_N$. We note that

$$s_1(N)-s_2(N)=\widetilde{s}_1(N-1)-\widetilde{s}_2(N-1)+1-\xi_N\geq\widetilde{s}_1(N-1)-\widetilde{s}_2(N-1),$$

so in this case $\min_{1\leq j\leq N}\{s_1(j)-s_2(j)\}=\min_{1\leq j\leq N-1}\{\widetilde{s}_1(j)-\widetilde{s}_2(j)\}$. Thus (2.7) gives

$$\begin{aligned}f_N(\tau_1,\dots,\tau_{N-1},1)&=\sum_{\xi_N\in\{0,1\}}\sum_{\xi_1,\dots,\xi_{N-1}\in\{0,1\}^{N-1}}\mathsf{b}^{1-\xi_N}\frac{\mathsf{b}^{\widetilde{s}_1(N-1)-\widetilde{s}_2(N-1)}}{(\mathsf{ab})^{\min_{1\leq j\leq N-1}\{\widetilde{s}_1(j)-\widetilde{s}_2(j)\}}}\\&=\sum_{\xi_N\in\{0,1\}}\mathsf{b}^{1-\xi_N}\sum_{\xi_1,\dots,\xi_{N-1}\in\{0,1\}^{N-1}}\frac{\mathsf{b}^{\widetilde{s}_1(N-1)-\widetilde{s}_2(N-1)}}{(\mathsf{ab})^{\min_{1\leq j\leq N-1}\{\widetilde{s}_1(j)-\widetilde{s}_2(j)\}}}\\&=(1+\mathsf{b})f_{N-1}(\tau_1,\dots,\tau_{N-1}).\end{aligned}$$

Finally, we verify that for a fixed $1\leq n\leq N-1$ we have

$$\begin{aligned}f_N(\tau_1,\dots,\tau_{n-1},1,0,\tau_{n+2},\dots,\tau_N)&=f_{N-1}(\tau_1,\dots,\tau_{n-1},1,\tau_{n+2},\dots,\tau_N)\\&\quad+f_{N-1}(\tau_1,\dots,\tau_{n-1},0,\tau_{n+2},\dots,\tau_N).\end{aligned}\tag{2.9}$$

Here for $k \leq n-1$ we let $\widetilde{s}_1(k) = s_1(k)$ and $\widetilde{s}_2(k) = s_2(k)$. For $n \leq k \leq N-1$ we set $\widetilde{s}_2(k) = s_2(n-1) + \sum_{j=n}^{k} \xi_{j+1}$, skipping over ξ_n. On the other hand, for $n \leq k \leq N-1$ we let $\widetilde{s}_1(k) = s_1(n-1) + (1-\xi_n) + \sum_{j=n+1}^{k} \tau_{j+1}$. (This is one place in the proof where $\widetilde{\boldsymbol{s}}_1$ depends on both $\boldsymbol{\tau}$ and $\boldsymbol{\xi}$.) Note that putting this choice of $\widetilde{\boldsymbol{s}}_1$ into expression (2.7) leads to the formula for $f_{N-1}(\tau_1, \ldots, \tau_{n-1}, 1-\xi_n, \tau_{n+2}, \ldots, \tau_N)$.

It is clear that

$$s_1(N) - s_2(N) = \left(\sum_{j=1}^{n-1} \tau_j + 1 + 0 + \sum_{j=n+2}^{N} \tau_j \right) - \left(\sum_{j=1}^{n-1} \xi_j + \xi_n + \sum_{j=n+1}^{N} \xi_j \right)$$
$$= \left(\sum_{j=1}^{n-1} \tau_j + (1-\xi_n) + \sum_{j=n+1}^{N-1} \tau_{j+1} \right) - \left(\sum_{j=1}^{n-1} \xi_j + \sum_{j=n}^{N-1} \xi_{j+1} \right) = \widetilde{s}_1(N-1) - \widetilde{s}_2(N-1). \tag{2.10}$$

The same calculation shows that $s_1(k) - s_2(k) = \widetilde{s}_1(k-1) - \widetilde{s}_2(k-1)$ for $k = n+1, \ldots, N$. Since $s_1(k) - s_2(k) = \widetilde{s}_1(k) - \widetilde{s}_2(k)$ for $0 \leq k \leq n-1$, and by the same rewrite as in (2.10) we get

$$s_1(n) - s_2(n) = 1 - \xi_n + s_1(n-1) - s_2(n-1) \geq s_1(n-1) - s_2(n-1),$$

we see that $s_1(n) - s_2(n)$ does not contribute to the minimum. This shows that the two minima are the same,

$$\min_{0 \leq k \leq N} \{s_1(k) - s_2(k)\} = \min_{0 \leq k \leq n-1} \{s_1(k) - s_2(k)\} \wedge \min_{n+1 \leq k \leq N} \{s_1(k) - s_2(k)\}$$
$$= \min_{0 \leq k \leq n-1} \{\widetilde{s}_1(k) - \widetilde{s}_2(k)\} \wedge \min_{n+1 \leq k \leq N} \{\widetilde{s}_1(k-1) - \widetilde{s}_2(k-1)\} = \min_{0 \leq k \leq N-1} \{\widetilde{s}_1(k) - \widetilde{s}_2(k)\}.$$

Therefore summing in (2.7) over $\xi_1, \ldots, \xi_{n-1}, \xi_{n+1}, \ldots, \xi_N \in \{0,1\}$ and isolating the sum over $\xi_n \in \{0,1\}$ we get

$$f_N(\tau_1, \ldots, \tau_{n-1}, 1, 0, \tau_{n+2}, \ldots, \tau_N) = \sum_{\xi_n \in \{0,1\}} f_{N-1}(\tau_1, \ldots, \tau_{n-1}, 1-\xi_n, \tau_{n+2}, \ldots, \tau_N),$$

which establishes (2.9).

Since $f_N(\tau_1, \ldots, \tau_N)$ and $p_N(\tau_1, \ldots, \tau_N)$ satisfy the same recursion with respect to N and the same initial conditions at $N = 1$, this ends the proof. □

Proof of Theorem 1.1. For $\boldsymbol{s}_1, \boldsymbol{s}_2 \in \mathcal{S}$, denote

$$g_N(\boldsymbol{s}_1, \boldsymbol{s}_2) = \frac{\mathsf{b}^{s_1(N) - s_2(N)}}{(\mathsf{ab})^{\min_{0 \leq j \leq N} \{s_1(j) - s_2(j)\}}}, \tag{2.11}$$

and let $\hat{g}_N(\boldsymbol{\tau}, \boldsymbol{\xi})$ denote the same expression treated as a function of $\boldsymbol{\tau}, \boldsymbol{\xi} \in \Omega^{(N)}$ under the bijection (2.6). (That is, we apply (2.11) to $\boldsymbol{s}_1$ which is the height function of $\boldsymbol{\tau}$ and to $\boldsymbol{s}_2$, which is the height function of $\boldsymbol{\xi}$.) In this notation, (2.7) becomes

$$f_N(\boldsymbol{\tau}) = \sum_{\boldsymbol{\xi} \in \Omega^{(N)}} \hat{g}_N(\boldsymbol{\tau}, \boldsymbol{\xi}). \tag{2.12}$$

Since $\mathbb{P}_{\mathrm{rw}}$ is a uniform law on $\mathcal{S} \times \mathcal{S}$, formula (1.3) can be written as

$$\mathbb{P}_{\mathrm{TLE}}(\boldsymbol{s}_1, \boldsymbol{s}_2) = \frac{1}{\mathsf{Z}_{\mathsf{a},\mathsf{b}}} g_N(\boldsymbol{s}_1, \boldsymbol{s}_2),$$

where $\mathsf{Z}_{\mathsf{a},\mathsf{b}}$ is the same normalizing constant as in (2.1). Indeed, the normalizing constant is

$$\sum_{\boldsymbol{s}_1,\boldsymbol{s}_2\in\mathcal{S}} g_N(\boldsymbol{s}_1,\boldsymbol{s}_2)=\sum_{\boldsymbol{\tau},\boldsymbol{\xi}\in\Omega_T}\hat{g}_N(\boldsymbol{\tau},\boldsymbol{\xi})=\sum_{\boldsymbol{\tau}\in\Omega_T} f_N(\boldsymbol{\tau})=\sum_{\boldsymbol{\tau}\in\Omega_T} p_N(\boldsymbol{\tau})=\mathsf{Z}_{\mathsf{a},\mathsf{b}},$$

where we used (2.12) and Lemma 2.1.

Thus, writing $\boldsymbol{\tau}=\boldsymbol{\tau}(\boldsymbol{s}_1)$ and $\boldsymbol{\xi}=\boldsymbol{\xi}(\boldsymbol{s}_2)$ for the inverse of bijection (2.6), we have

$$\begin{aligned}\sum_{\boldsymbol{s}_2\in\mathcal{S}}\mathbb{P}_{\mathrm{TLE}}(\boldsymbol{s}_1,\boldsymbol{s}_2)&=\frac{1}{\mathsf{Z}_{\mathsf{a},\mathsf{b}}}\sum_{\boldsymbol{s}_2\in\mathcal{S}} g_N(\boldsymbol{s}_1,\boldsymbol{s}_2)=\frac{1}{\mathsf{Z}_{\mathsf{a},\mathsf{b}}}\sum_{\boldsymbol{\xi}\in\Omega}\hat{g}_N(\boldsymbol{\tau}(\boldsymbol{s}_1),\boldsymbol{\xi})\\&=\frac{1}{\mathsf{Z}_{\mathsf{a},\mathsf{b}}}f_N(\boldsymbol{\tau}(\boldsymbol{s}_1))=\frac{1}{\mathsf{Z}_{\mathsf{a},\mathsf{b}}}p_N(\boldsymbol{\tau}(\boldsymbol{s}_1))=\mathbb{P}_{\mathrm{TASEP}}(\boldsymbol{\tau}(\boldsymbol{s}_1))=\mathbb{P}_{\mathrm{H}}(\boldsymbol{s}_1),\end{aligned}$$

proving (1.5). □

3 Applications

In this section, we use Theorem 1.1 to refine and extend existing results concerning the asymptotic behavior of the height function in steady state. Our two theorems draw inspiration and build upon the foundational works of [12] where non-Gaussian fluctuations were identified, and [16] which described the rate function for large deviations.

3.1 Stationary measure of the conjectural KPZ fixed point on a segment

Barraquand and Le Doussal [2] introduced process $B+X$, where B,X are independent processes, B is the Brownian motion of variance $1/2$, and the law of X is given by the Radon-Nikodym derivative

$$\frac{d\mathbb{P}_X}{d\mathbb{P}_B}(\omega)=\frac{1}{\mathfrak{K}(\mathsf{u},\mathsf{v})}\exp\left((\mathsf{u}+\mathsf{v})\min_{0\le x\le 1}\omega(x)-\mathsf{v}\omega(1)\right),\quad \omega\in C[0,1]. \tag{3.1}$$

They proposed this process as the stationary measure of the conjectural KPZ fixed point on the interval $[0,1]$ with boundary parameters $\mathsf{u},\mathsf{v}\in\mathbb{R}$, and predicted that the process $B+X$ should arise as a scaling limit of stationary measures of all models in the KPZ universality class on an interval.

This prediction is supported by the limit theorem established in [7], which demonstrated that the process $B+X$ describes the asymptotic behavior of the fluctuations of the height function of ASEP with parameters that vary with the size of the system according to formula (3.2) under condition $\mathsf{u}+\mathsf{v}>0$. In this paper, we extend the above limit theorem to all real values of u and v, removing the positivity constraint $\mathsf{u}+\mathsf{v}>0$ for TASEP. This, to our knowledge, constitutes the first confirmation of the prediction in [2] that encompasses the entire range of real parameters u and v. A recent preprint [27] uses our result in combination with analytic techniques to extend Theorem 3.1 to the general ASEP.

Recall that the height function is defined by (1.2).

Theorem 3.1. *Consider a sequence of TASEPs indexed by N on the segments $\{1,\dots,N\}$ for all $N\in\mathbb{N}$, with parameters $\alpha=\alpha(N)\to 1/2$, $\beta=\beta(N)\to 1/2$ as $N\to\infty$ at the rates given by relation* (1.1) *with*

$$\mathsf{a}=\mathsf{a}_N=e^{-\mathsf{u}/\sqrt{N}},\quad \mathsf{b}=\mathsf{b}_N=e^{-\mathsf{v}/\sqrt{N}}. \tag{3.2}$$

Then

$$\left\{\frac{1}{\sqrt{N}}\left(2H_N(\lfloor xN\rfloor)-\lfloor Nx\rfloor\right)\right\}_{x\in[0,1]}\Rightarrow\{B_x+X_x\}_{x\in[0,1]}\ \text{as } N\to\infty,$$

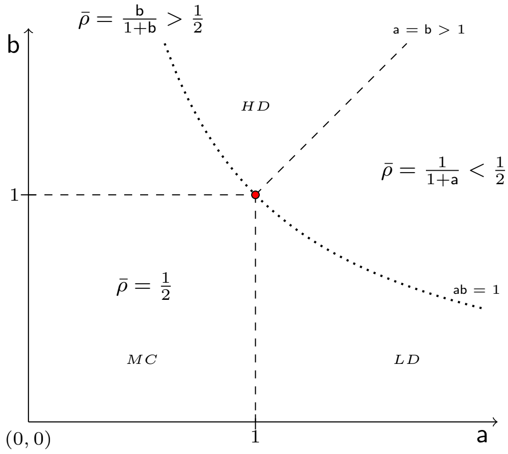

Figure 2: For fixed a, b, limiting particle density $\bar\rho := \lim_{N\to\infty} \frac{1}{N}(\tau_1 + \cdots + \tau_N)$ varies by region of the phase diagram for the open TASEP. These are the maximal current region, marked as MC, the low density region, marked as LD, and the high density region, marked as HD. Hyperbola $ab = 1$ separates the *fan region* from the *shock region* $ab > 1$. (These regions were identified in [11, Fig. 3].) Parameters a_N, b_N of the TASEP in Theorem 3.1 vary with system size N and converge to the *triple point* $(a, b) = (1, 1)$, where the regions MC, LD, and HD meet.

where the convergence is in Skorokhod's space $D[0,1]$ of càdlàg functions, processes B, X are independent processes on $[0,1]$ with continuous trajectories, B is a Brownian motion of variance $1/2$, and the law of X is given by (3.1).

We remark that a linear interpolation of the height function similar to the one that appears in Theorem 3.4 leads to the same conclusion under weak convergence in the space $C[0,1]$ with the supremum norm.

Proof. Consider the coordinate processes $\boldsymbol{S}_1, \boldsymbol{S}_2$ defined on the probability space $(\mathcal{S}^{(N)} \times \mathcal{S}^{(N)}, \mathbb{P}_{\mathrm{TLE}})$ by (1.4). We shall determine the limit of the process

$$\boldsymbol{W}_1^{(N)} := \frac{1}{\sqrt{N}} \left\{2S_1(\lfloor xN \rfloor) - \lfloor xN \rfloor\right\}_{x\in[0,1]}. \tag{3.3}$$

We will establish a more general claim that under measure $\mathbb{P}_{\mathrm{TLE}}$, we have joint convergence of the pair of processes:

$$\frac{1}{\sqrt{N}} \left(\{S_1(\lfloor xN \rfloor) + S_2(\lfloor xN \rfloor) - \lfloor xN \rfloor\}_{x\in[0,1]}, \{S_1(\lfloor xN \rfloor) - S_2(\lfloor xN \rfloor)\}_{x\in[0,1]}\right)$$
$$\Rightarrow \left(\{B_x\}_{x\in[0,1]}, \{X_x\}_{x\in[0,1]}\right) \text{ as } N \to \infty, \tag{3.4}$$

where B, X are independent processes from the conclusion of the theorem. Noting that the process (3.3) is the sum of the processes on the left hand side of (3.4), this will end the proof.

Fix a bounded continuous function $\Phi : D([0,1]; \mathbb{R}^2) \to \mathbb{R}$ and write

$$\boldsymbol{W}_\pm^{(N)} := \tfrac12(\boldsymbol{W}_1^{(N)} \pm \boldsymbol{W}_2^{(N)})$$

for the two processes on the left hand side of (3.4), where $\boldsymbol{W}_2^{(N)}$ is defined as in (3.3) with $\boldsymbol{S}_2$ in place of $\boldsymbol{S}_1$. Our goal is to prove that

$$\lim_{N\to\infty} \mathbb{E}_{\mathrm{TLE}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)})\right] = \mathbb{E}[\Phi(B, X)]. \tag{3.5}$$

Following the approach in [9, Proposition 1.3], we use Donsker's theorem. Since $\mathrm{Var}_{\mathrm{rw}}(S_1(N)) = N/4$, by Donsker's theorem, under probability measure $\mathbb{P}_{\mathrm{rw}}$, we have

$$(\boldsymbol{W}_1^{(N)}, \boldsymbol{W}_2^{(N)}) \Rightarrow (W_1, W_2), \quad \text{as } N \to \infty,$$

where W_1, W_2 are independent Wiener processes and convergence is in $D([0,1];\mathbb{R}^2)$. Thus

$$(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) \Rightarrow (B, B') := (W_1 + W_2, W_1 - W_2)/2,$$

where B, B' are independent Brownian motions of variance $1/2$. Using the Radon-Nikodym density (1.3) and (3.2), we get

$$\begin{aligned}
\mathbb{E}_{\mathrm{TLE}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)})\right] &= \frac{1}{\mathsf{Z}_{\mathsf{u},\mathsf{v}}(N)} \mathbb{E}_{\mathrm{rw}}\Big[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) \\
&\quad \times \exp\left(\frac{\mathsf{u}+\mathsf{v}}{\sqrt{N}} \min_{0\le j\le N}(S_1(j) - S_2(j)) - \frac{\mathsf{v}}{\sqrt{N}}(S_1(N) - S_2(N))\right)\Big] \\
&= \frac{1}{\mathsf{Z}_{\mathsf{u},\mathsf{v}}(N)} \mathbb{E}_{\mathrm{rw}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) e^{(\mathsf{u}+\mathsf{v})\min_{0\le x\le 1} \boldsymbol{W}_-^{(N)}(x) - \mathsf{v}\boldsymbol{W}_-^{(N)}(1)}\right] \\
&= \frac{1}{\mathsf{Z}_{\mathsf{u},\mathsf{v}}(N)} \mathbb{E}_{\mathrm{rw}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) \mathcal{E}(\boldsymbol{W}_-^{(N)})\right],
\end{aligned}$$

where

$$\mathcal{E}(f) = e^{(\mathsf{u}+\mathsf{v})\min_{0\le x\le 1} f(x) - \mathsf{v} f(1)}.$$

Noting that Φ is bounded and by (A.1) and Remark A.2

$$\sup_N \mathbb{E}_{\mathrm{rw}}\left[\mathcal{E}^2(\boldsymbol{W}_-^{(N)})\right] < \infty,$$

we see that the sequence of real valued random variables

$$\left\{\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) \mathcal{E}(\boldsymbol{W}_-^{(N)})\right\}_{N=1,2,\ldots}$$

is uniformly integrable with respect to $\mathbb{P}_{\mathrm{rw}}$. Uniform integrability and weak convergence imply convergence of expectations ([5, Theorem 3.5]), so Donsker's theorem implies that

$$\lim_{N\to\infty} \mathbb{E}_{\mathrm{rw}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) \mathcal{E}(\boldsymbol{W}_-^{(N)})\right] = \mathbb{E}_{B,B'}\left[\Phi(B, B')\mathcal{E}(B')\right],$$

where $\mathbb{E}_{B,B'}$ denotes integration with respect to the law of (B, B') on $C[0,1] \times C[0,1]$. Using this with $\Phi \equiv 1$, we see that the normalizing constants also converge,

$$\lim_{N\to\infty} \mathsf{Z}_{\mathsf{u},\mathsf{v}}(N) = \lim_{N\to\infty} \mathbb{E}_{\mathrm{rw}}\left[\mathcal{E}(\boldsymbol{W}_-^{(N)})\right] = \mathbb{E}_{B,B'}\left[\mathcal{E}(B')\right] = \mathfrak{K}(\mathsf{u},\mathsf{v}).$$

Thus

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}_{\mathrm{TLE}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)})\right] &= \frac{\lim_{N\to\infty} \mathbb{E}_{\mathrm{rw}}\left[\Phi(\boldsymbol{W}_+^{(N)}, \boldsymbol{W}_-^{(N)}) \mathcal{E}(\boldsymbol{W}_-^{(N)})\right]}{\lim_{N\to\infty} \mathsf{Z}_{\mathsf{u},\mathsf{v}}(N)} \\
= \frac{1}{\mathfrak{K}(\mathsf{u},\mathsf{v})} \mathbb{E}_{B,B'}\left[\Phi(B, B')\mathcal{E}(B')\right] &= \int_{C[0,1]} \left(\int_{C[0,1]} \Phi(b, b') \frac{\mathcal{E}(b')}{\mathfrak{K}(\mathsf{u},\mathsf{v})} P_B(db')\right) P_B(db) \\
&= \int_{C[0,1]} \left(\int_{C[0,1]} \Phi(b, x) P_X(dx)\right) P_B(db) = \mathbb{E}_{B,X}\left[\Phi(B, X)\right],
\end{aligned}$$

where we used independence of the Brownian motions B, B' and (3.1). (Here, $\mathbb{E}_{B,X}$ denotes integration with respect to the (product) law of (B, X) on $C[0,1] \times C[0,1]$.) This establishes (3.5) and ends the proof. □

3.2 Large deviations

Large deviations for the height function process $\{H_N(\lfloor Nx\rfloor)\}_{x\in[0,1]}$ of ASEP have been discussed in Refs. [17] and [16], with the height function interpreted as particle density profile. (There are also nice expositions in [18, Section 5], [19, Section 16].) In this section, we use Theorem 1.1 to deduce large deviations for the height function of the TASEP directly from Mogulskii's theorem [10, Theorem 5.1.2]. In Section 4 we show how to recover formulas discovered in [16], and we determine the additive normalization.

Let $(\mathbb{X},d)$ be a complete separable metric space. Consider a sequence of probability spaces $(\Omega^{(N)},\mathbb{P}^{(N)})$ and a family of random variables $X_N:\Omega^{(N)}\to\mathbb{X}$, $N=1,2,\dots$. A standard statement of the large deviation principle in Varadhan's sense involves a family of Borel subsets A of $\mathbb{X}$, their interiors $\mathrm{int(A)}$ and closures $\mathrm{cl(A)}$ and specifies asymptotics of probabilities in terms of a rate function I by the following upper/lower bounds:

$$-\inf_{x\in\mathrm{int(A)}} I(x)\le \liminf_{N\to\infty}\frac{1}{N}\log\mathbb{P}^{(N)}(X_N\in A)$$
$$\le \limsup_{N\to\infty}\frac{1}{N}\log\mathbb{P}^{(N)}(X_N\in A)\le -\inf_{x\in\mathrm{cl(A)}} I(x).$$

It will be more convenient to use an equivalent definition which we now recall; see [10, Theorem 4.4.13]. (Compare also [22, Definitions 1.1.1 and 1.2.2].)

Definition 3.2. *Let $(\mathbb{X},d)$ be a complete separable metric space. The sequence $\{X_N\}$ satisfies the* large deviation principle *(LDP), if there exists a lower semicontinuous function $I:\mathbb{X}\to[0,\infty]$, called the* rate function*, such that*

(i) for every bounded continuous function $\Phi:\mathbb{X}\to\mathbb{R}$,

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}_N\left[\exp(N\Phi(X_N))\right]=\sup_{x\in\mathbb{X}}\{\Phi(x)-I(x)\},\tag{3.6}$$

where $\mathbb{E}_N[\cdot]=\int_{\Omega^{(N)}}(\cdot)d\mathbb{P}_N$ denotes the expected value.

(ii) I has compact level sets, $I^{-1}[0,w]$ is a compact subset of $\mathbb{X}$ for every $w\ge 0$.

To prove the LDP for the height function, we consider the sequence $\{X_N\}$ of $\mathbb{X}=C([0,1],\mathbb{R}^2)$-valued random variables defined on probability spaces $\Omega^{(N)}=(\mathcal{S}^{(N)}\times\mathcal{S}^{(N)},\mathbb{P}^{(N)}_{\mathrm{TLE}})$ obtained by linear interpolation between the points $\left(\frac{k}{N},\frac{S_1(k)}{N}\right)$, $k=0,\dots,N$, in the first component and the points $\left(\frac{k}{N},\frac{S_2(k)}{N}\right)$, $k=0,\dots,N$, in the second component. By Theorem 1.1, the first component of X_N then has the same law as the continuous interpolation of the height function (1.2) based on the points $\left(\frac{k}{N},\frac{H_N(k)}{N}\right)$, $k=0,\dots,N$.

To introduce the rate function, we need additional notation. Let

$$h(x)=\begin{cases}x\log x+(1-x)\log(1-x) & 0\le x\le 1,\\ \infty & x<0 \text{ or } x>1.\end{cases}$$

Denote by $\mathcal{AC}_0$ the set of absolutely continuous functions $f\in C[0,1]$ such that $f(0)=0$. Let

$$K(\mathsf{a},\mathsf{b})=\log(\bar\rho(1-\bar\rho)),\tag{3.7}$$

where

$$\bar\rho=\bar\rho(\mathsf{a},\mathsf{b})=\begin{cases}\frac{1}{1+\mathsf{a}} & \mathsf{a}>1,\mathsf{a}<\mathsf{b},\\ \frac{1}{2} & \mathsf{a}\le 1,\mathsf{b}\le 1,\\ \frac{\mathsf{b}}{1+\mathsf{b}} & \mathsf{b}>1,\mathsf{b}>\mathsf{a},\end{cases}\tag{3.8}$$

denotes the limiting particle density, as indicated on the phase diagram in Fig. 2. (Notation $\log(\bar\rho(1-\bar\rho))$ was introduced in [16].)

Theorem 3.3. *If* $\mathsf{a},\mathsf{b}>0$*, then the sequence* $\{X_N\}$ *satisfies the large deviation principle with respect to the probability measures* $\mathbb{P}^{(N)}_{\rm TLE}$ *with the rate function*

$$I(f_1,f_2)=\int_0^1 \left(h(f_1'(x))+h(f_2'(x))\right)dx \\ +\log(\mathsf{ab})\min_{0\le x\le 1}(f_1(x)-f_2(x))-(f_1(1)-f_2(1))\log\mathsf{b}-K(\mathsf{a},\mathsf{b}) \tag{3.9}$$

if functions $f_1,f_2\in\mathcal{AC}_0$*; we let* $I=\infty$ *for all other* $f_1,f_2\in C[0,1]$*.*

Since $h=\infty$ outside of $[0,1]$, it is clear that expression (3.9) can only be finite for $f_1,f_2\in\mathcal{AC}_0$ with the derivatives in $[0,1]$ for almost all x. In particular, as a consequence of the Arzelà-Ascoli theorem, $I(\cdot)$ is lower semicontinuous and has compact level sets.

Proof. By a theorem of Mogulskii, see [10, Theorem 5.1.2], X_N satisfies the LDP with respect to the law $\mathbb{P}_{\rm rw}$ of two independent Bernoulli$(1/2)$ random walk paths. The rate function for two independent components with Bernoulli$(1/2)$ increments becomes

$$I_{\rm rw}(f_1,f_2)=\log 4+\int_0^1 \left(h(f_1'(x))+h(f_2'(x))\right)dx=\int_0^1\left(h(f_1'(x)|\tfrac12)+h(f_2'(x)|\tfrac12)\right)dx,$$

where $f_1,f_2\in\mathcal{AC}_0$, see [10, Exercise 2.2.23], and

$$h(x|y)=x\log\left(\frac{x}{y}\right)+(1-x)\log\left(\frac{1-x}{1-y}\right),\quad x,y\in(0,1). \tag{3.10}$$

In order to use (3.6), we need to fix a bounded continuous function $\Phi: C([0,1],\mathbb{R}^2)\to\mathbb{R}$ and compute

$$\lim_{N\to\infty}\frac1N\log\mathbb{E}_{\rm TLE}\left[e^{N\Phi(X_N)}\right]=\lim_{N\to\infty}\frac1N\log\mathbb{E}_{\rm rw}\left[e^{N\Phi(X_N)}g(X_N)^N\right], \tag{3.11}$$

where g is given by (2.11). To proceed, we introduce a version of g that acts on pairs of continuous functions $f_j\in\mathcal{AC}_0$ by

$$\widetilde g(f_1,f_2):=\frac{\mathsf{b}^{f_1(1)-f_2(1)}}{(\mathsf{ab})^{\min_{0\le x\le1}(f_1(x)-f_2(x))}}.$$

It is clear that $\widetilde g(NX_N)=g(\boldsymbol S_1,\boldsymbol S_2)$, where $\boldsymbol S_k=(0,S_k(1),\dots,S_k(N))$ are the sums from (1.4) and that $\widetilde g(NX_N)=\widetilde g(X_N)^N=\exp(N\log\widetilde g(X_N))$. To compute the limit (3.11), we use Varadhan's lemma with respect to $\mathbb{P}_{\rm rw}$. According to Varadhan's lemma [10, Theorem 4.3.1], if $\Psi: C([0,1],\mathbb{R}^2)\to\mathbb{R}$ is a continuous function such that

$$\limsup_{N\to\infty}\frac1N\log\mathbb{E}_{\rm rw}\left[e^{\gamma N\Psi(X_N)}\right]<\infty \text{ for some } \gamma>1, \tag{3.12}$$

then

$$\lim_{N\to\infty}\frac1N\log\mathbb{E}_{\rm rw}\left[e^{N\Psi(X_N)}\right]=\sup_{f_1,f_2}\{\Psi(f_1,f_2)-I_{\rm rw}(f_1,f_2)\}.$$

We apply this to a continuous but unbounded function $\Psi: C([0,1],\mathbb{R}^2)\to\mathbb{R}$ given by

$$\Psi(f_1,f_2)=\Phi(f_1,f_2)+\log\widetilde g(f_1,f_2) \\ =\Phi(f_1,f_2)+\log(\mathsf{b})(f_1(1)-f_2(1))-\log(\mathsf{ab})\min_{0\le x\le1}(f_1(x)-f_2(x)).$$

To verify that (3.12) holds with $\gamma = 2$, we note that by the Cauchy-Schwartz inequality

$$\mathbb{E}_{\rm rw}\left[e^{2N\Psi(X_N)}\right] = \mathbb{E}_{\rm rw}\left[e^{2N\Phi(X_N)} g^2(\boldsymbol{S}_1,\boldsymbol{S}_2)\right] \leq e^{N\|\Phi\|_\infty}\mathbb{E}_{\rm rw}\left[g^2(\boldsymbol{S}_1,\boldsymbol{S}_2)\right]$$
$$\leq e^{N\|\Phi\|_\infty}\left(\mathbb{E}_{\rm rw}\left[(\mathsf{ab})^{-4\min_{0\leq j\leq N}\{S_1(j)-S_2(j)\}}\right]\right)^{1/2}\left(\mathbb{E}_{\rm rw}\left[\mathsf{b}^{4(S_1(N)-S_2(N))}\right]\right)^{1/2}.$$

Inequality (A.1), see Remark A.2, shows that (3.12) holds. From Varadhan's lemma [10, Theorem 4.3.1] we get

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}_{\rm rw}\left[e^{N\Phi(X_N)}\widetilde{g}(X_N)^N\right] = \sup_{f_1,f_2\in C[0,1]}\{\Phi(f_1,f_2)+\log\widetilde{g}(f_1,f_2)-I_{\rm rw}(f_1,f_2)\}$$
$$= \sup_{f_1,f_2\in C[0,1]}\Big\{\Phi(f_1,f_2)-\big(I_{\rm rw}(f_1,f_2)$$
$$+\log(\mathsf{ab})\min_{0\leq x\leq 1}(f_1(x)-f_2(x))-(f_1(1)-f_2(1))\log\mathsf{b}\big)\Big\}.$$

Using this with $\Phi\equiv 0$ we see that the normalizing constants in (1.3) satisfy

$$\lim_{N\to\infty}\frac{1}{N}\log\mathfrak{C}_{\mathsf{a},\mathsf{b}}(N) = -K_0(\mathsf{a},\mathsf{b}),$$

where $K_0(\mathsf{a},\mathsf{b})$ is given by the same variational expression

$$K_0(\mathsf{a},\mathsf{b}) = \inf_{f_1,f_2\in\mathcal{AC}_0}\Big\{\int_0^1\big(h(f_1'(x)|\tfrac{1}{2})+h(f_2'(x)|\tfrac{1}{2})\big)\,dx$$
$$+\log(\mathsf{ab})\min_{0\leq x\leq 1}(f_1(x)-f_2(x))-(f_1(1)-f_2(1))\log\mathsf{b}\Big\}. \quad (3.13)$$

Thus

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}_{\rm TLE}\left[e^{N\Phi(X_N)}\right] = \lim_{N\to\infty}\frac{1}{N}\log\mathbb{E}_{\rm rw}\left[e^{N\Phi(X_N)}\widetilde{g}(X_N)^N\right]-\lim_{N\to\infty}\frac{1}{N}\log\mathfrak{C}_{\mathsf{a},\mathsf{b}}(N)$$
$$= \sup_{f_1,f_2\in C[0,1]}\Big\{\Phi(f_1,f_2)$$
$$-\left(I_{\rm rw}(f_1,f_2)+\log(\mathsf{ab})\min_{0\leq x\leq 1}(f_1(x)-f_2(x))-(f_1(1)-f_2(1))\log\mathsf{b}-K_0(\mathsf{a},\mathsf{b})\right)\Big\}.$$

This proves LDP with the rate function that depends on $K_0(\mathsf{a},\mathsf{b})$. To conclude the proof, it remains to verify that expression (3.13) for $K_0(\mathsf{a},\mathsf{b})$ simplifies to the expression $K(\mathsf{a},\mathsf{b}) = \log(\bar\rho(1-\bar\rho))$ given in (3.7). We postpone this part of the proof until Section 4, where as part of the proof of Proposition 4.1, we will obtain formula (3.7) for $\mathsf{ab}\geq 1$, and then as part of the proof of Proposition 4.6 we will establish that (3.7) holds also for $\mathsf{ab}\leq 1$. □

It is clear that the continuous linear interpolation $\widetilde{H}_N$ of the step function forming the height function $\{H_N(\lfloor Nx\rfloor)\}_{x\in[0,1]}$ has the same law as the first component of the vector $NX_N(x)$. By contraction principle (see e.g. [10, Theorem 4.2.1] or [22, Theorem 1.3.2]), this implies the following LDP.

Theorem 3.4. *If $\mathsf{a},\mathsf{b}>0$ then the sequence of linear interpolations $\{\frac{1}{N}\widetilde{H}_N\}$ of the height function of a TASEP on $\{1,\dots,N\}$ satisfies the large deviation principle with the rate function*

$$\mathcal{I}(f) = \inf_{g\in\mathcal{AC}_0}\Big\{\int_0^1(h(f'(x))+h(g'(x)))dx$$
$$+\log(\mathsf{ab})\min_{0\leq x\leq 1}(f(x)-g(x))-\log(\mathsf{b})(f(1)-g(1))\Big\}-\log(\bar\rho(1-\bar\rho)) \quad (3.14)$$

if $f\in\mathcal{AC}_0$, and $\mathcal{I}(f)=\infty$ otherwise. Here $\bar\rho$ is given by (3.8).

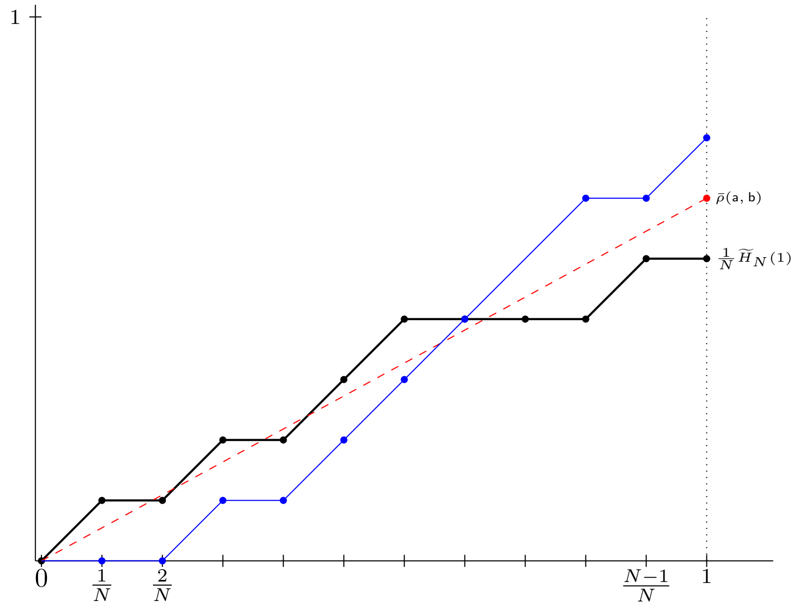

Figure 3: Continuous interpolation X_N is a pair of piecewise-linear lines. The first component $\frac{1}{N}\widetilde{H}_N$ of X_N is marked as the thick black line and the second component of X_N marked in blue. Note that $\frac{1}{N}\widetilde{H}_N(1) = \frac{1}{N}\sum_{j=1}^N \tau_j \to \bar{\rho}$, see Fig. 2, except on the coexistence line with $\mathsf{a} = \mathsf{b} > 1$. (The dashed line represents the most likely trajectory of the random curve $x \mapsto \frac{1}{N}\widetilde{H}_N(x)$ for large N.)

Proof. To obtain (3.14), we apply the contraction principle to the first coordinate mapping, with (3.9) applied to $f_1 = f$, $f_2 = g$. Anticipating subsequent developments we wrote (3.7) for $K_0(\mathsf{a},\mathsf{b})$. □

A version of this result for a more general ASEP appears in [16], with the rate function rewritten in different and more explicit forms for $\mathsf{ab} \le 1$ and $\mathsf{ab} > 1$, which are discussed in Section 4. The main novelty in Theorem 3.4 is that its proof and the expression for the rate function do not distinguish between the shock and the fan region. (This has been anticipated in [4, Formula (3.56)].) On the other hand, additional nontrivial work is needed to recover the formulas that appear in [16].

We remark that Ref. [4, Section 3.6] uses a two layer representation [21] to obtain a version of Theorem 3.4 for the case $\mathsf{a} = \mathsf{b} = 0$, which is not covered by our results. Ref. [4] also discusses the corresponding version of Proposition 4.2.

4 Comparison with previous large deviation results

In this section we discuss previous LDP results for ASEP, specialized to the case of TASEP. With some additional work these results can be obtained from Theorem 3.4, and the derivations identify constant $K_0(\mathsf{a},\mathsf{b})$ given by the variational formula (3.13) in the proof as a simpler expression (3.7) based on the phase diagram.

Converted to our notation, the LDP in Ref. [16, (1.11)] gives the following rate function for the shock region of TASEP:

Proposition 4.1. *If* $\mathsf{ab} \geq 1$ *then the rate function* (3.14) *is*

$$\mathcal{I}(f) = \min_{0\leq y\leq 1} \Big\{ \int_0^y \Big(f'(x)\log\tfrac{\mathsf{a}f'(x)}{1+\mathsf{a}} + (1-f'(x))\log\tfrac{1-f'(x)}{1+\mathsf{a}}\Big)\,dx \\ + \int_y^1 \Big(f'(x)\log\tfrac{f'(x)}{1+\mathsf{b}} + (1-f'(x))\log\tfrac{(1-f'(x))\mathsf{b}}{1+\mathsf{b}}\Big)\,dx\Big\} - K(\mathsf{a},\mathsf{b}), \quad (4.1)$$

where

$$K(\mathsf{a},\mathsf{b}) = \log\min\left\{\frac{\mathsf{a}}{(1+\mathsf{a})^2}, \frac{\mathsf{b}}{(1+\mathsf{b})^2}\right\} = \log\frac{\mathsf{a}\vee\mathsf{b}}{(1+\mathsf{a}\vee\mathsf{b})^2}. \quad (4.2)$$

Although this result is known, we provide a separate proof based on Theorem 3.4. This allows us to complete part of the postponed proof of Theorem 3.4, where we need to show that $K_0(\mathsf{a},\mathsf{b}) = K(\mathsf{a},\mathsf{b}) = \log\bar\rho(1-\bar\rho)$ for $\mathsf{ab}\geq 1$. To accomplish this goal, we use (3.14) with $K_0(\mathsf{a},\mathsf{b})$ given by (3.13). (Then the fact that $K_0(\mathsf{a},\mathsf{b}) = \log\bar\rho(1-\bar\rho)$ for $\mathsf{ab}\geq 1$ will follow from (4.2).)

Proof of Proposition 4.1. With $\log(\mathsf{ab})\geq 0$, we write (3.14) as

$$\mathcal{I}(f) = -K_0(\mathsf{a},\mathsf{b}) + \int_0^1 h(f'(x))dx \\ + \inf_g \min_{0\leq y\leq 1}\Big\{\int_0^1 h(g'(x))dx + \log(\mathsf{ab})(f(y)-g(y)) - \log(\mathsf{b})(f(1)-g(1))\Big\} \\ = -K_0(\mathsf{a},\mathsf{b}) + \int_0^1 h(f'(x))dx + \min_{0\leq y\leq 1}\Big\{f(y)\log(\mathsf{ab}) - f(1)\log\mathsf{b} \\ + \inf_{g_1\in\mathcal{AC}_0[0,y]}\Big\{\int_0^y h(g_1'(x))dx - \log(\mathsf{ab})g_1(y)\Big\} \\ + \inf_{g_2\in\mathcal{AC}[y,1]:g_2(y)=g_1(y)}\Big\{\int_y^1 h(g_2'(x))dx + \log(\mathsf{b})g_2(1)\Big\}\Big\},$$

where we split g into $g_1 = g|_{[0,y]}$, $g_2 = g|_{[y,1]}$ for $0<y<1$, with (omitted) minor changes for $y=0$ or $y=1$. Since $\inf_g \int_a^b h(g'(x))dx$ is attained on linear functions, denoting by $F = g_1(y) = g_2(y)$ and $G = g_2(1)$ the values of g_2 at the endpoints of interval $[y,1]$, the optimal (linear) functions are $\widetilde{g}_1(x) = Fx/y$ and

$$\widetilde{g}_2(x) = \frac{(G-F)(x-y)}{1-y} + F.$$

Optimizing over all possible choices of $F\leq G$ we get

$$\inf_{g_1}\int_0^y h(g_1'(x))dx - \log(\mathsf{ab})g_1(y) + \inf_{g_2}\int_y^1 h(g_2'(x))dx + \log(\mathsf{b})g_2(1) \\ = \min_{F\in[0,y]}\min_{G\in[F,F+1-y]}\{yh(F/y) - \log(\mathsf{ab})F + (1-y)h(\tfrac{G-F}{1-y}) + G\log\mathsf{b}\} \\ = y\log\frac{1}{1+\mathsf{a}} + (1-y)\log\left(\frac{\mathsf{b}}{1+\mathsf{b}}\right),$$

as the minimum over F,G is attained at

$$F = \frac{\mathsf{a}y}{1+\mathsf{a}}, \quad G = F + \frac{1-y}{1+\mathsf{b}}.$$

We get

$$\mathcal{I}(f) = -K_0(\mathsf{a},\mathsf{b}) + \int_0^1 h(f'(x))dx + \min_{y\in[0,1]} \Big\{ f(y)\log \mathsf{a} + y\log\tfrac{1}{1+\mathsf{a}} + (1-y)\log\tfrac{\mathsf{b}}{1+\mathsf{b}} - (f(1)-f(y))\log \mathsf{b}. \Big\}$$
$$= -K_0(\mathsf{a},\mathsf{b}) + \min_{y\in[0,1]} \Big\{ \int_0^y h(f'(x))dx + f(y)\log\mathsf{a} + y\log\tfrac{1}{1+\mathsf{a}} + \int_y^1 h(f'(x))dx + (1-y)\log\tfrac{\mathsf{b}}{1+\mathsf{b}} - (f(1)-f(y))\log\mathsf{b} \Big\}.$$

To end the proof, we note that the two integrals under the minimum in (4.1) match the two integrals in the expression above:

$$\int_0^y \Big(f'(x)\log\tfrac{\mathsf{a}f'(x)}{1+\mathsf{a}} + (1-f'(x))\log\tfrac{1-f'(x)}{1+\mathsf{a}} \Big) dx = \int_0^y h(f'(x))dx + f(y)\log\mathsf{a} + y\log\tfrac{1}{1+\mathsf{a}}$$

and

$$\int_y^1 \Big(f'(x)\log\tfrac{f'(x)}{1+\mathsf{b}} + (1-f'(x))\log\tfrac{(1-f'(x))\mathsf{b}}{1+\mathsf{b}} \Big) dx = \int_y^1 h(f'(x))dx + (f(1)-f(y))\log\tfrac{1}{1+\mathsf{b}} + (1-y)\log\tfrac{\mathsf{b}}{1+\mathsf{b}} - (f(1)-f(y))\log\tfrac{\mathsf{b}}{1+\mathsf{b}}$$
$$= \int_y^1 h(f'(x))dx + (1-y)\log\tfrac{\mathsf{b}}{1+\mathsf{b}} - (f(1)-f(y))\log\mathsf{b}.$$

The additive normalization constant $K_0(\mathsf{a},\mathsf{b})$ can now be determined from the condition that $\inf_f \mathcal{I}(f)=0$. To do so, we repeat the previous calculation again. Denoting by $F=f(y)$ and $G=f(1)$, we switch the order of the infima, and use the extremal property of linear functions again:

$$K_0(\mathsf{a},\mathsf{b}) = \min_{y\in[0,1]} \inf_f \Big\{ \int_0^y h(f'(x))dx + \int_y^1 h(f'(x))dx + f(y)\log\mathsf{a} + y\log\tfrac{1}{1+\mathsf{a}} + (1-y)\log\tfrac{\mathsf{b}}{1+\mathsf{b}} - (f(1)-f(y))\log\mathsf{b}. \Big\}$$
$$= \min_{y\in[0,1]} \inf_{F\in[0,y], G\in[F,F+1-y]} \Big\{ yh(F/y) + (1-y)h(\tfrac{G-F}{1-y}) + F\log\mathsf{a} + y\log\tfrac{1}{1+\mathsf{a}} + (1-y)\log\tfrac{\mathsf{b}}{1+\mathsf{b}} - (G-F)\log\mathsf{b} \Big\}.$$

The infimum is attained at $F=\frac{y}{1+\mathsf{a}}$, $G = F + \frac{\mathsf{b}}{1+\mathsf{b}}(1-y)$. Therefore,

$$K_0(\mathsf{a},\mathsf{b}) = \min_{y\in[0,1]} \Big\{ y\log\frac{\mathsf{a}}{(1+\mathsf{a})^2}\log\Big(\frac{\mathsf{a}}{\mathsf{b}}\Big) - 2y\log(1+\mathsf{a}) + 2(y-1)\log(1+\mathsf{b}) + \log(\mathsf{b}) \Big\}$$
$$= \min_{y\in[0,1]} \Big\{ y\log\Big(\frac{\mathsf{a}}{(1+\mathsf{a})^2}\Big) - (1-y)\log\Big(\frac{\mathsf{b}}{(1+\mathsf{b})^2}\Big) \Big\}$$

and (4.2) follows. This establishes (3.7) for $\mathsf{ab}\geq 1$. □

For the fan region of TASEP, the LDP in [16, (1.7), (3.3), (3.6)] gives the rate function which we recalculated as follows.

Proposition 4.2. *If* $\mathsf{ab} < 1$ *then for* $f \in \mathcal{AC}_0$ *with* $0 \le f' \le 1$ *the rate function* (3.14) *is*

$$\mathcal{I}(f) = \int_0^1 h(f'(x))dx + \int_0^1 (\widetilde{f}'(x) \log G_*(x) + (1 - \widetilde{f}') \log(1 - G_*(x)))dx - K(\mathsf{a}, \mathsf{b}), \quad (4.3)$$

where $G_*(x) = \left(\widetilde{f}'(x) \vee \frac{\mathsf{a}}{1+\mathsf{a}}\right) \wedge \frac{1}{1+\mathsf{b}}$, *and* $\widetilde{f}$ *is the convex envelope of* f, *i.e., the largest convex function below* f. *The normalizing constant is*

$$K(\mathsf{a}, \mathsf{b}) = \sup_{\frac{\mathsf{b}}{1+\mathsf{b}} \le \rho \le \frac{1}{1+\mathsf{a}}} \log \rho(1 - \rho) = \begin{cases} \log \frac{\mathsf{a}}{(1+\mathsf{a})^2} & \mathsf{a} > 1, \\ -2 \log 2 & \mathsf{a}, \mathsf{b} \le 1, \\ \log \frac{\mathsf{b}}{(1+\mathsf{b})^2} & \mathsf{b} > 1. \end{cases} \quad (4.4)$$

We verify that this result follows from Theorem 3.4. As previously, to avoid circular reasoning we use (3.14) with $K_0(\mathsf{a}, \mathsf{b})$ given by (3.13), without identification $K_0(\mathsf{a}, \mathsf{b}) = \log \bar{\rho}(1 - \bar{\rho})$. The identification will follow for $\mathsf{ab} < 1$ once we establish (4.4) in the proof of Proposition 4.6 below.

The proof is more substantial and requires additional lemmas, so we put it as a separate section.

4.1 Proof of Proposition 4.2

In the proof, $f \in \mathcal{AC}_0$ with $0 \le f' \le 1$ is fixed and $\widetilde{f}$ is its convex envelope. Since $\widetilde{f}$ is convex, $\widetilde{f}'$ is nondecreasing, and for definiteness, we take $\widetilde{f}'$ right-continuous. We write (3.14) as $\mathcal{I}(f) = \int_0^1 h(f'(x))dx - K(\mathsf{a}, \mathsf{b}) + \inf_g J^*(f, g)$, where

$$J^*(f, g) = \int_0^1 h(g'(x))dx + \log(\mathsf{ab}) \min_{0 \le x \le 1} (f(x) - g(x)) - \log(\mathsf{b})(f(1) - g(1)). \quad (4.5)$$

Similarly, we write (4.3) as $\mathcal{I}(f) = \int_0^1 h(f'(x))dx - K(\mathsf{a}, \mathsf{b}) + J_*(\widetilde{f}, G_*)$, where

$$J_*(f, G) = \int_0^1 [f'(x) \log G(x) + (1 - f'(x)) \log(1 - G(x))] \, dx. \quad (4.6)$$

(One can verify that $J_*(\widetilde{f}, G_*) = J_*(f, G_*)$, see the proof of [16, (A.5)], and it is the latter expression that appears in the rate function [16, (3.3)].) We want to prove that

$$J_*(\widetilde{f}, G_*) = \inf_{g \in \mathcal{AC}_0} J^*(f, g). \quad (4.7)$$

We first verify that

Lemma 4.3.

$$\inf_{g \in \mathcal{AC}_0} J^*(f, g) = \inf_{g \in \mathcal{AC}_0} J^*(\widetilde{f}, g). \quad (4.8)$$

Proof. For any g the inequality $J^*(f, g) \le J^*(\tilde{f}, g)$ is trivial because $f \ge \widetilde{f}$, so $\min(f - g) \ge \min(\widetilde{f} - g)$, $\log(\mathsf{ab}) < 0$, and $f(1) = \widetilde{f}(1)$.

For the converse inequality, consider the set $U = \{f > \widetilde{f}\}$. This set is open and therefore is a disjoint union of open intervals, say $U = \bigcup J_k$. On each of these intervals, $\widetilde{f}$ is linear. Define a function G as follows: $G(x) = g'(x)$ for $x \notin U$ and

$$G(x) = \frac{g(w_k) - g(u_k)}{w_k - u_k} \text{ for } x \in J_k = (u_k, w_k).$$

Then $\int_{J_k} G = g(w_k) - g(u_k)$. Define $\widetilde{g}(x) = \int_0^x G$. We have

$$\widetilde{g}(x) = \int_{[0,x] \setminus U} g'(x) + \sum_k \mathbb{I}_{\{w_k < x\}}(g(w_k) - g(u_k)) + \mathbb{I}_{\{x \in J_k\}}(x - u_k) \frac{g(w_k) - g(u_k)}{w_k - u_k},$$

thus $\widetilde{g}$ is linear on each J_k and $\widetilde{g} = g$ on $[0,1] \setminus U$. Then $\widetilde{f} - \widetilde{g}$ is also linear on J_k, and

$$\begin{aligned}\min_{J_k}(\widetilde{f} - \widetilde{g}) &= \min(\widetilde{f}(u_k) - \widetilde{g}(u_k), \widetilde{f}(w_k) - \widetilde{g}(w_k)) \\ &= \min(f(u_k) - g(u_k), f(w_k) - g(w_k)) \geq \min_{[0,1]}(f - g).\end{aligned}$$

Thus, $\min_{[0,1]}(\widetilde{f} - \widetilde{g}) \geq \min_{[0,1]}(f - g)$. From convexity of h, we have

$$\int_{J_k} h(g') \geq |J_k| h\Big(\frac{1}{|J_k|}\int_{J_k} g'\Big) = |J_k| h\Big(\frac{g(w_k) - g(u_k)}{w_k - u_k}\Big) = \int_{J_k} h(\widetilde{g}').$$

Thus

$$\int_0^1 h(g') \geq \int_0^1 h(\widetilde{g}').$$

Therefore, we obtain $J^*(f,g) \geq J^*(\widetilde{f}, \widetilde{g})$, completing the proof of (4.8). □

Clearly, $\frac{\mathsf{a}}{1+\mathsf{a}} < \frac{1}{1+\mathsf{b}}$. For clarity, we write G_* in expanded form

$$G_*(x) = \begin{cases} \frac{\mathsf{a}}{1+\mathsf{a}} & \widetilde{f}'(x) < \frac{\mathsf{a}}{1+\mathsf{a}}, \\ \widetilde{f}'(x) & \frac{\mathsf{a}}{1+\mathsf{a}} \leq \widetilde{f}'(x) < \frac{1}{1+\mathsf{b}}, \\ \frac{1}{1+\mathsf{b}} & \widetilde{f}'(x) \geq \frac{1}{1+\mathsf{b}}, \end{cases} \tag{4.9}$$

and we note that G_* is nondecreasing and right-continuous.

We can now relate the functionals J_* and J^*.

Lemma 4.4. *Let* $g_*(x) = \int_0^x G_*$. *Then*

$$J^*(\widetilde{f}, g_*) = J_*(\widetilde{f}, G_*). \tag{4.10}$$

Proof. With the convention $\inf \emptyset = 1$ and $\sup \emptyset = 0$, let

$$x_1 = \inf\left\{x \geq 0 : \widetilde{f}'(x) \geq \frac{\mathsf{a}}{1+\mathsf{a}}\right\}, \quad x_2 = \sup\left\{x \leq 1 : \widetilde{f}'(x) < \frac{1}{1+\mathsf{b}}\right\} \tag{4.11}$$

be the largest interval $[x_1, x_2)$ on which $G_* = \widetilde{f}'$. Under this convention, we have $x_1 = x_2 = 1$ when $\widetilde{f}' < \mathsf{a}/(1+\mathsf{a})$ on $[0,1]$ and $x_1 = x_2 = 0$ when $\widetilde{f}' \geq 1/(1+\mathsf{b})$ on [0,1]. It is also possible that G_* jumps from its lowest to its largest value at some x_* in which case we have $x_1 = x_2 = x_*$.

From (4.11) we see that $\widetilde{f}' < G_* = g_*'$ on $[0, x_1)$, $\widetilde{f}' = G_* = g_*'$ on $[x_1, x_2)$, and $\widetilde{f}' > G_* = g_*'$ on $(x_2, 1]$. Then $\widetilde{f} - g_*$ decreases on $[0, x_1)$, is constant on $[x_1, x_2]$, and increases on $(x_2, 1]$. Thus the minimal value of $\widetilde{f} - g_*$ is attained on the entire interval $[x_1, x_2]$.

Identity (4.10) now follows by direct calculation. We have

$$\begin{aligned} J_*(\widetilde{f}, G_*) &= \int_0^{x_1}\left(\widetilde{f}'\log\frac{\mathsf{a}}{1+\mathsf{a}} + (1-\widetilde{f}')\log\frac{1}{1+\mathsf{a}}\right) + \int_{x_1}^{x_2} h(\widetilde{f}') \\ &\quad + \int_{x_2}^1\left(\widetilde{f}'\log\frac{1}{1+\mathsf{b}} + (1-\widetilde{f}')\log\frac{\mathsf{b}}{1+\mathsf{b}}\right) \\ &= x_1\log\frac{1}{1+\mathsf{a}} + \widetilde{f}(x_1)\log\mathsf{a} + \int_{x_1}^{x_2} h(\widetilde{f}') + (1-x_2)\log\frac{\mathsf{b}}{1+\mathsf{b}} - (\widetilde{f}(1) - \widetilde{f}(x_2))\log\mathsf{b}. \end{aligned} \tag{4.12}$$

On the other hand, since the minimum of $\widetilde{f} - g_*$ is attained at all points of $[x_1, x_2]$, we get

$$\log(\mathsf{ab}) \min_{[0,1]}(\widetilde{f} - g_*) - (\widetilde{f}(1) - g_*(1)) \log \mathsf{b}$$
$$= (\widetilde{f}(x_1) - g_*(x_1)) \log \mathsf{a} - (\widetilde{f}(1) - \widetilde{f}(x_2)) \log \mathsf{b} + (g_*(1) - g_*(x_2)) \log \mathsf{b}. \quad (4.13)$$

Since $g_*(x_1) = x_1 \frac{\mathsf{a}}{1+\mathsf{a}}$ and $g_*(1) - g_*(x_2) = (1 - x_2) \frac{1}{1+\mathsf{b}}$,

$$\int_0^1 h(g_*') = \int_0^{x_1} h\left(\frac{\mathsf{a}}{1+\mathsf{a}}\right) + \int_{x_1}^{x_2} h(\widetilde{f}') + \int_{x_2}^1 h\left(\frac{1}{1+\mathsf{b}}\right)$$
$$= x_1 \log \frac{1}{1+\mathsf{a}} + x_1 \frac{\mathsf{a}}{1+\mathsf{a}} \log \mathsf{a} + \int_{x_1}^{x_2} h(\widetilde{f}') - (1-x_2)\frac{1}{1+\mathsf{b}} \log \mathsf{b} + (1-x_2) \log \frac{\mathsf{b}}{1+\mathsf{b}}$$
$$= x_1 \log \frac{1}{1+\mathsf{a}} + g_*(x_1) \log \mathsf{a} + \int_{x_1}^{x_2} h(\widetilde{f}') - (g_*(1) - g_*(x_2)) \log \mathsf{b} + (1-x_2) \log \frac{\mathsf{b}}{1+\mathsf{b}}. \quad (4.14)$$

Combining (4.13) and (4.14), we get

$$J^*(\widetilde{f}, g_*) = x_1 \log \frac{1}{1+\mathsf{a}} + \widetilde{f}(x_1) \log \mathsf{a} + \int_{x_1}^{x_2} h(\widetilde{f}') + (1-x_2) \log \frac{\mathsf{b}}{1+\mathsf{b}} - (\widetilde{f}(1) - \widetilde{f}(x_2)) \log \mathsf{b}$$

which ends the proof by (4.12). □

With (4.8) and (4.10) at hand, to complete the proof of (4.7), we need to show that for any $g \in \mathcal{AC}_0$ the following inequality holds:

$$J^*(\widetilde{f}, g) \geq J^*(\widetilde{f}, g_*), \quad (4.15)$$

hence, by (4.8), the infimum $\inf_g J^*(f, g) = \inf_g J^*(\widetilde{f}, g)$ is attained at g_*.

To prove (4.15), we fix $g \in \mathcal{AC}_0$ and consider the function

$$\mathsf{F}(\tau) = J^*(\widetilde{f}, \tau g + (1-\tau) g_*), \qquad \tau \in [0, 1].$$

Note that this function is convex because the functional $J^*(\widetilde{f}, \cdot)$ is convex. We claim that the right-derivative $\partial_+ \mathsf{F}(0) \geq 0$. When this is proved, the convexity implies that F is increasing on $[0, 1]$, and therefore

$$J^*(\widetilde{f}, g) = \mathsf{F}(1) \geq \mathsf{F}(0) = J^*(\widetilde{f}, g_*).$$

We need the following technical lemma.

Lemma 4.5. *Let ϕ and ψ be continuous functions on $[0, 1]$. Let $A = \{x \in [0, 1] \colon \phi(x) = \min_{[0,1]} \phi\}$. Then*

$$\min_{[0,1]}(\phi - \tau\psi) = \min_{[0,1]} \phi - \tau \cdot \max_A \psi + o(\tau) \quad \text{as } \tau \to 0^+.$$

Proof. First, the set A is closed and therefore ψ reaches the maximum on it, say, at z. Then for $\tau \geq 0$ we have

$$\min_{[0,1]}(\phi - \tau\psi) \leq \phi(z) - \tau\psi(z) = \min_{[0,1]} \phi - \tau \cdot \max_A \psi.$$

To prove the converse estimate, let us fix any $\varepsilon_1 > 0$ and using the continuity of ψ find $\delta > 0$ such that $\max_{A_\delta} \psi \leq \max_A \psi + \varepsilon_1$, where $A_\delta = \{y \in [0, 1] \colon \operatorname{dist}(y, A) < \delta\}$. Then use continuity of ϕ to find $\varepsilon_2 > 0$ such that $\min_{[0,1]\setminus A_\delta} \phi > \min_A \phi + \varepsilon_2$. For a small positive τ we have

$$\min_{[0,1]\setminus A_\delta} (\phi - \tau\psi) \geq \min_A \phi + \varepsilon_2 - \tau \max_{[0,1]} \psi \geq \min_{[0,1]} \phi - \tau \max_A \psi;$$
$$\min_{A_\delta} (\phi - \tau\psi) \geq \min_A \phi - \tau \max_{A_\delta} \psi \geq \min_{[0,1]} \phi - \tau \max_A \psi - \tau\varepsilon_1.$$

Combining these inequalities, we obtain

$$\liminf_{\tau\to 0+} \frac{\min_{[0,1]}(\phi - \tau\psi) - \min_{[0,1]} \phi}{\tau} \geq -\max_A \psi - \varepsilon_1.$$

Tending ε_1 to zero, we finish the proof. □

Proof of Proposition 4.2. We apply Lemma 4.5 with $\phi = \widetilde{f} - g_*$ and $\psi = g - g_*$ to obtain

$$\min_{[0,1]}(\widetilde{f}-(\tau g+(1-\tau)g_*)) = \min_{[0,1]}(\widetilde{f}-g_*-\tau(g-g_*)) = \min_{[0,1]}(\widetilde{f}-g_*)-\tau \max_A(g-g_*)+o(\tau),\ \tau \to 0^+,$$

where $A = \{x \in [0,1]\colon \widetilde{f}(x) - g_*(x) = \min_{[0,1]}(\widetilde{f} - g_*)\} = [x_1, x_2]$. We use this formula to calculate the right derivative of F at 0:

$$\begin{aligned}
\partial_+\mathsf{F}(0) &= \int_0^1 (g' - g_*') \log\left(\frac{g_*'}{1-g_*'}\right) - \log(\mathsf{ab}) \max_{[x_1,x_2]}(g - g_*) + (g(1) - g_*(1)) \log \mathsf{b}\\
&= \int_0^1 (g' - g_*') \log\left(\frac{\mathsf{b}g_*'}{1-g_*'}\right) - \log(\mathsf{ab}) \max_{[x_1,x_2]}(g - g_*)\\
&= \int_0^{x_1} (g' - g_*') \log\left(\frac{\mathsf{b}\frac{\mathsf{a}}{\mathsf{a}+1}}{1-\frac{\mathsf{a}}{\mathsf{a}+1}}\right) + \int_{x_1}^{x_2} (g' - g_*') \log\left(\frac{\mathsf{b}g_*'}{1-g_*'}\right) + \int_{x_2}^1 (g' - g_*') \log\left(\frac{\mathsf{b}\frac{1}{\mathsf{b}+1}}{1-\frac{1}{\mathsf{b}+1}}\right)\\
&\quad - \log(\mathsf{ab}) \max_{[x_1,x_2]}(g - g_*)\\
&= \log(\mathsf{ab}) \int_0^{x_1} (g' - g_*') + \int_{x_1}^{x_2} (g' - g_*') \log\left(\frac{\mathsf{b}g_*'}{1-g_*'}\right) - \log(\mathsf{ab}) \max_{[x_1,x_2]}(g - g_*)\\
&= \int_{x_1}^{x_2} (g' - g_*') \log\left(\frac{\mathsf{b}g_*'}{1-g_*'}\right) - \log(\mathsf{ab})\Big(\max_{[x_1,x_2]}(g - g_*) - \big(g(x_1) - g_*(x_1)\big)\Big).
\end{aligned} \tag{4.16}$$

If $x_1 = x_2$, this gives $\partial_+\mathsf{F}(0) = 0$. If $x_1 < x_2$, we proceed as follows.

Consider the function $\Phi = \log\left(\frac{\mathsf{b}g_*'}{1-g_*'}\right)$ on $[x_1, x_2]$. It is a nondecreasing right-continuous function with values $\Phi(x_1) \geq \log(\mathsf{ab})$ and $\Phi(x_2) \leq 0$. Write $\Phi(x) - \Phi(x_1) = \mu([x_1, x])$, where μ is a non-negative measure of total mass $\mu([x_1, x_2]) \in [0, -\log(\mathsf{ab})]$ on Borel subsets of $[x_1, x_2]$.

Write $\psi = g' - g_*'$ and $\Psi(x) = \int_{x_1}^x \psi = g(x) - g_*(x) - (g(x_1) - g_*(x_1))$.

Since $\Phi(x) - \Phi(x_1) = \mu([x_1, x])$ for $x \in [x_1, x_2]$ and $\Psi(x_1) = 0$, Fubini' theorem gives

$$\begin{aligned}
L := \int_{x_1}^{x_2} \psi\Phi = \int_{x_1}^{x_2} \psi(x) \left(\Phi(x_1) + \int_{[x_1,x_2]} 1_{t\leq x} d\mu(t)\right) dx\\
= \Psi(x_2)\Phi(x_1) + \int_{[x_1,x_2]} \int_t^{x_2} \psi(x) dx d\mu(t)\\
= \Psi(x_2)\Phi(x_1) + \int_{[x_1,x_2]} (\Psi(x_2) - \Psi(t)) d\mu(t) = \Psi(x_2)\Phi(x_2) - \int_{[x_1,x_2]} \Psi(t) d\mu(t).
\end{aligned}$$

Since $\Phi(x_2) \leq 0$, we get $\Psi(x_2)\Phi(x_2) \geq \Phi(x_2) \max_{[x_1,x_2]} \Psi$, therefore

$$L = \Psi(x_2)\Phi(x_2) - \int_{[x_1,x_2]} \Psi(z)\ d\mu(z) \geq (\Phi(x_2) - \mu([x_1, x_2])) \max_{[x_1,x_2]} \Psi.$$

Thus
$$\int_{x_1}^{x_2} \psi\Phi \geq \Phi(x_1) \max_{[x_1,x_2]} \Psi.$$
Returning back to (4.16), we see that since $\max_{[x_1,x_2]} \Psi \geq \Psi(x_1) = 0$, and $\Phi(x_1) \geq \log(\mathsf{ab})$, we have
$$\partial_+ \mathsf{F}(0) = \int_{x_1}^{x_2} \psi\Phi - \log(\mathsf{ab}) \max_{[x_1,x_2]} \Psi \geq (\Phi(x_1) - \log(\mathsf{ab})) \max_{[x_1,x_2]} \Psi \geq 0.$$
This proves (4.15). To prove (4.7), we combine (4.8), (4.10), and (4.15). This concludes the proof of Proposition 4.2. ☐

4.2 Large deviations for the mean particle density

The mean particle density is
$$\frac{1}{N}\sum_{j=1}^{N} \tau_j = \tfrac{1}{N} H_N(N) = \tfrac{1}{N} \widetilde{H}_N(1).$$
The following proposition, recalculated from [16, formula (3.12)], gives explicit formula for the rate function of the mean particle density in the fan region of TASEP. An equivalent result with a different proof appeared in [8, Theorem 7].

Proposition 4.6. *If* $\mathsf{ab} \leq 1$*, then the mean particle density* $\frac{1}{N} H_N(N)$ *satisfies the large deviation principle with the rate function*
$$\mathsf{I}(r) = -K(\mathsf{a},\mathsf{b}) + \begin{cases} h(r|\frac{1}{1+\mathsf{a}}) + \log \frac{\mathsf{a}}{(1+\mathsf{a})^2} & 0 \leq r < \frac{\mathsf{a}}{1+\mathsf{a}}, \\ 2h(r|\frac{1}{2}) + \log\frac{1}{4} & \frac{\mathsf{a}}{1+\mathsf{a}} \leq r \leq \frac{1}{1+\mathsf{b}}, \\ h(r|\frac{\mathsf{b}}{1+\mathsf{b}}) + \log \frac{\mathsf{b}}{(1+\mathsf{b})^2} & \frac{1}{1+\mathsf{b}} < r \leq 1, \end{cases} \tag{4.17}$$
with $K(\mathsf{a},\mathsf{b})$ *given by* (4.4) *and entropy* $h(\cdot\,|\,\cdot)$ *given by* (3.10)*.*

Although the result is known, we re-derive formula (4.17) from (3.9) for the special case of TASEP, as the argument establishes (4.4), and hence we will conclude the derivation of formula (3.7) for $\mathsf{ab} \leq 1$.

Proof of Proposition 4.6. We use (3.9) and contraction principle. To avoid circular reasoning, we use (3.9) with $K_0(\mathsf{a},\mathsf{b})$ given by (3.13). (In fact, we leave $K_0(\mathsf{a},\mathsf{b})$ as a free parameter to be determined at the end of the proof.)

Through the proof, we fix $r \in [0,1]$. We write $\mathsf{I}(r)$ as the infimum over $m \in [0,1]$ and over all functions $f_1, f_2 \in \mathcal{AC}_0$ such that $f_1(1) = r$, $f_2(1) = m$. The first step is to show that optimal f_1, f_2 are linear. To do so we note that since $\mathsf{ab} \leq 1$, we have
$$\log(\mathsf{ab}) \min_{y\in[0,1]} \{f_1(y) - f_2(y)\} \geq \log(\mathsf{ab}) \min\{0, (r-m)\}$$
with equality on linear functions. Therefore, the expression $\log(\mathsf{ab}) \min_y \{f_1(y) - f_2(y)\}$ can only decrease if we replace f_1, f_2 with a pair of linear functions $f(x) = rx$ and $g(x) = mx$. In view of convexity of h, this replacement also decreases the integral in (3.9). This shows that the optimal functions f_1, f_2 are indeed linear, $f(x) = rx$ and $g(x) = mx$. We get
$$\mathsf{I}(r) = -K_0(\mathsf{a},\mathsf{b}) + h(r) + \min_{m\in[0,1]} \left\{ h(m) + \max_{y\in[0,1]} \{y \log(\mathsf{ab})(r-m) - \log \mathsf{b}(r-m)\} \right\}.$$
The maximum over $y \in [0,1]$ is attained at the end points of $[0,1]$ and since $\log(\mathsf{ab}) \leq 0$, it is either $(r-m)\log \mathsf{a}$ or $(m-r)\log \mathsf{b}$ depending on whether $m \geq r$ or $m < r$. (Recall that $r \in [0,1]$ is fixed.)

In the first case, the infimum over $m \geq r$ is attained at

$$m = \begin{cases} \frac{\mathsf{a}}{1+\mathsf{a}} & \text{if } r < \frac{\mathsf{a}}{1+\mathsf{a}}, \\ r & \text{if } r \geq \frac{\mathsf{a}}{1+\mathsf{a}}, \end{cases}$$

and gives

$$\mathsf{I}_1(r) = -K_0(\mathsf{a},\mathsf{b}) + \begin{cases} h(r) + r\log \mathsf{a} - \log(1+\mathsf{a}) & r < \frac{\mathsf{a}}{1+\mathsf{a}}, \\ 2h(r) & r \geq \frac{\mathsf{a}}{1+\mathsf{a}}, \end{cases}$$

with $K_0(\mathsf{a},\mathsf{b})$ given by (3.13). Note that

$$h(r) + r\log \mathsf{a} - \log(1+\mathsf{a}) = h(r|\tfrac{1}{1+\mathsf{a}}) + \log \tfrac{\mathsf{a}}{(1+\mathsf{a})^2}.$$

In the second case, the minimum over $m \leq r$ is attained at

$$m = \begin{cases} \frac{1}{1+\mathsf{b}} & \text{if } r > \frac{1}{1+\mathsf{b}}, \\ r & \text{if } r \leq \frac{1}{1+\mathsf{b}}, \end{cases}$$

and gives

$$\mathsf{I}_2(r) = -K_0(\mathsf{a},\mathsf{b}) + \begin{cases} h(r) - r\log(\mathsf{b}) + \log \frac{\mathsf{b}}{\mathsf{b}+1} & r > \frac{1}{1+\mathsf{b}}, \\ 2h(r) & r \leq \frac{1}{1+\mathsf{b}}. \end{cases}$$

Note that

$$h(r) - r\log(\mathsf{b}) + \log \tfrac{\mathsf{b}}{\mathsf{b}+1} = h(r|\tfrac{\mathsf{b}}{1+\mathsf{b}}) + \log \tfrac{\mathsf{b}}{(1+\mathsf{b})^2}.$$

Also, note that $h(r) = h(r|\frac{1}{2}) - \log 2$. Since we are interested in overall minimum over all $m \in [0,1]$, up to the additive normalizing constant $K_0(\mathsf{a},\mathsf{b})$, the resulting rate function is

$$\mathsf{I}(r) = \min\{\mathsf{I}_1(r), \mathsf{I}_2(r)\} = -K_0(\mathsf{a},\mathsf{b}) + \begin{cases} h(r|\frac{1}{1+\mathsf{a}}) + \log \frac{\mathsf{a}}{(1+\mathsf{a})^2} & 0 \leq r < \frac{\mathsf{a}}{1+\mathsf{a}}, \\ h(r|\frac{\mathsf{b}}{1+\mathsf{b}}) + \log \frac{\mathsf{b}}{(1+\mathsf{b})^2} & \frac{1}{1+\mathsf{b}} < r \leq 1, \\ 2h(r|2) - \log 4 & \text{ortherwise}, \end{cases}$$

in agreement with (4.17).

The additive normalization constant $K_0(\mathsf{a},\mathsf{b})$ can now be determined from the condition that $\inf_{r\in[0,1]} \mathsf{I}(r) = 0$. This gives (4.4) as follows:

$$K_0(\mathsf{a},\mathsf{b}) = \inf_{r\in[0,1]} \begin{cases} h(r|\frac{1}{1+\mathsf{a}}) + \log \frac{\mathsf{a}}{(1+\mathsf{a})^2} & 0 \leq r < \frac{\mathsf{a}}{1+\mathsf{a}}, \\ h(r|\frac{\mathsf{b}}{1+\mathsf{b}}) + \log \frac{\mathsf{b}}{(1+\mathsf{b})^2} & \frac{1}{1+\mathsf{b}} < r \leq 1, \\ 2h(r|\frac{1}{2}) - \log 4 & \frac{\mathsf{a}}{1+\mathsf{a}} \leq r \leq \frac{1}{1+\mathsf{b}}. \end{cases}$$

Thus $K_0(\mathsf{a},\mathsf{b}) = \log(\bar\rho(1-\bar\rho))$, matching (4.4). We note that this establishes (3.7) for $\mathsf{ab} \leq 1$. □

The LDP for the mean particle density in the shock region follows from Proposition 4.1 by the contraction principle. The following proposition, recalculated from [16, (B.8)], gives an explicit formula for the rate function.

Proposition 4.7. *If $\mathsf{ab} > 1$, then the mean particle density $\frac{1}{N}H_N(N)$ satisfies the large deviation principle with the rate function*

$$\mathsf{I}(r) = \begin{cases} h(r|\frac{1}{1+\mathsf{a}}) + \log \frac{\mathsf{a}}{(1+\mathsf{a})^2} - K(\mathsf{a},\mathsf{b}) & 0 \leq r \leq \frac{1}{1+\mathsf{b}}, \\ r\log \frac{\mathsf{a}}{\mathsf{b}} + \log \frac{\mathsf{b}}{(1+\mathsf{a})(1+\mathsf{b})} - K(\mathsf{a},\mathsf{b}) & \frac{1}{1+\mathsf{b}} \leq r \leq \frac{\mathsf{a}}{1+\mathsf{a}}, \\ h(r|\frac{\mathsf{b}}{1+\mathsf{b}}) + \log \frac{\mathsf{b}}{(1+\mathsf{b})^2} - K(\mathsf{a},\mathsf{b}) & \frac{\mathsf{a}}{1+\mathsf{a}} \leq r \leq 1, \end{cases} \tag{4.18}$$

with $\mathsf{I} = \infty$ for $r \notin [0,1]$.

Proof. The proof of this formula appears in [16, Section 3.6 and Appendix B, Case 2] and is omitted. □

We remark that on the coexistence line $\mathsf{a} = \mathsf{b} > 1$, the rate function is zero on the entire interval $[1/(1+\mathsf{a}), \mathsf{a}/(1+\mathsf{a})]$. This is consistent with [26, Theorem 1.6] which implies that mean particle density $\frac{1}{N} H_N(N)$ converges in distribution to the uniform law on this interval. (Shocks on the coexistence line for open ASEP were also described in [15, 20, 25].)

A Integrability lemma

The Lévy-Ottaviani maximal inequality and tail integration give the following bound:

Lemma A.1. *For* $\mathsf{c} > 0$*, we have*

$$\mathbb{E}_{\mathrm{rw}}\left[\mathsf{c}^{-\min_{0\le j\le N}(S_1(j)-S_2(j))}\right] \le 1 + 2\,\mathbb{E}_{\mathrm{rw}}\left[\mathsf{c}^{|S_1(N)-S_2(N)|}\right], \tag{A.1}$$

where $\mathbb{E}_{\mathrm{rw}}$ *is expectation with respect to the law of the two independent Bernoulli*$(1/2)$ *random walks.*

Proof. Recall that for sums of independent symmetric random variables $\{S_1(j) - S_2(j)\}$, the Lévy-Ottaviani maximal inequality says that

$$\mathbb{P}_{\mathrm{rw}}\left(\max_{1\le j\le N}|S_1(j)-S_2(j)| > t\right) \le 2\mathbb{P}_{\mathrm{rw}}\left(|S_1(N)-S_2(N)| > t\right), \quad t \ge 0.$$

Since

$$0 \le -\min_{0\le j\le N}(S_1(j)-S_2(j)) \le \max_{1\le j\le N}|S_1(j)-S_2(j)|,$$

we see that if $\mathsf{c} \le 1$ then the left hand side of (A.1) is bounded by 1. And if $\mathsf{c} > 1$, then by the above bound and tail integration, we have

$$\begin{aligned}
\mathbb{E}_{\mathrm{rw}}\left[\mathsf{c}^{-\min_{0\le j\le N}(S_1(j)-S_2(j))}\right] &\le \mathbb{E}_{\mathrm{rw}}\left[\mathsf{c}^{\max_{1\le j\le N}|S_1(j)-S_2(j)|}\right] \\
&= 1 + \log(\mathsf{c})\int_0^\infty \mathsf{c}^t\,\mathbb{P}_{\mathrm{rw}}\left(\max_{1\le j\le N}|S_1(j)-S_2(j)| > t\right)dt \\
&\le 1 + 2\log(\mathsf{c})\int_0^\infty \mathsf{c}^t\,\mathbb{P}_{\mathrm{rw}}\left(|S_1(N)-S_2(N)| > t\right)dt \\
&< 2\left(1 + \log(\mathsf{c})\int_0^\infty \mathsf{c}^t\,\mathbb{P}_{\mathrm{rw}}\left(|S_1(N)-S_2(N)| > t\right)dt\right) \\
&= 2\,\mathbb{E}_{\mathrm{rw}}\left[\mathsf{c}^{|S_1(N)-S_2(N)|}\right].
\end{aligned}$$

□

Remark A.2. Note that elementary inequalities $e^{|x|} \le e^x + e^{-x}$ and $1 + \cosh(x) \le 2e^{x^2/2}$ together with independence give

$$\begin{aligned}
\mathbb{E}_{\mathrm{rw}}\left[e^{\lambda|S_1(N)-S_2(N)|}\right] &\le \mathbb{E}_{\mathrm{rw}}\left[e^{\lambda(S_1(N)-S_2(N))}\right] + \mathbb{E}_{\mathrm{rw}}\left[e^{\lambda(S_2(N)-S_1(N))}\right] \\
&= 2\left(\frac{1+\cosh(\lambda)}{2}\right)^N \le 2e^{\lambda^2 N/2}.
\end{aligned}$$

In particular, using this with $\lambda = 2/\sqrt{N}$ we get a bound

$$\sup_N \mathbb{E}_{\mathrm{rw}}\left[\mathcal{E}^2(\boldsymbol{W}_-^{(N)})\right] < \infty$$

that is used in the proof of Theorem 3.1, and using this bound with $\lambda = -8\log(\mathsf{ab})$ and $\lambda = 8\log\mathsf{b}$, we get a bound

$$\sup_N \frac{1}{N}\log\mathbb{E}_{\rm rw}\left[e^{4N\log\tilde{g}(X_N)}\right] = \sup_N \frac{1}{N}\log\mathbb{E}_{\rm rw}\left[e^{4\log g(\boldsymbol{S}_1,\boldsymbol{S}_2)}\right]$$
$$\le \sup_N \frac{1}{2N}\log\mathbb{E}_{\rm rw}\left[(\mathsf{ab})^{-8\min_{0\le j\le N}\{S_1(j)-S_2(j)\}}\right] + \sup_N \frac{1}{2N}\log\mathbb{E}_{\rm rw}\left[\mathsf{b}^{8(S_1(N)-S_2(N))}\right]$$
$$\le \sup_N \frac{1}{2N}\log\left(1+2\mathbb{E}_{\rm rw}\left[\mathsf{c}^{|S_1(N)-S_2(N)|}\right]\right) + \sup_N \frac{1}{2N}\log\mathbb{E}_{\rm rw}\left[\mathsf{b}^{8(S_1(N)-S_2(N))}\right] < \infty,$$

which is used in the proof of Theorem 3.3.

References

[1] Guillaume Barraquand, Ivan Corwin, and Zongrui Yang, *Stationary measures for integrable polymers on a strip*, Invent. Math. **237** (2024), no. 3, 1567–1641, https://arxiv.org/pdf/2306.05983. MR4777093

[2] Guillaume Barraquand and Pierre Le Doussal, *Steady state of the KPZ equation on an interval and Liouville quantum mechanics*, Europhysics Letters **137** (2022), no. 6, 61003, ArXiv preprint with Supplementary material: https://arxiv.org/abs/2105.15178.

[3] Guillaume Barraquand and Pierre Le Doussal, *Stationary measures of the KPZ equation on an interval from Enaud-Derrida's matrix product ansatz representation*, J. Phys. A **56** (2023), no. 14, Paper No. 144003, 14 pp. MR4562516

[4] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Stochastic interacting particle systems out of equilibrium*, J. Stat. Mech. Theory Exp. (2007), no. 7, P07014, 35 pp. MR2335695

[5] Patrick Billingsley, *Convergence of probability measures*, second ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., New York, 1999, A Wiley-Interscience Publication. MR1700749

[6] Richard Brak, Sylvie Corteel, John Essam, Robert Parviainen, and Andrew Rechnitzer, *A combinatorial derivation of the PASEP stationary state*, Electron. J. Combin. **13** (2006), no. 1, Research Paper 108, 23 pp. MR2274323

[7] Włodek Bryc, Yizao Wang, and Jacek Wesołowski, *From the asymmetric simple exclusion processes to the stationary measures of the KPZ fixed point on an interval*, Ann. Inst. Henri Poincaré Probab. Stat. **59** (2023), no. 4, 2257–2284, https://arxiv.org/abs/2202.11869. MR4663522

[8] Włodek Bryc and Jacek Wesołowski, *Asymmetric simple exclusion process with open boundaries and quadratic harnesses*, J. Stat. Phys. **167** (2017), no. 2, 383–415, http://arxiv.org/abs/1511.01163. MR3626634

[9] Włodzimierz Bryc and Yizao Wang, *Fluctuations of random Motzkin paths II*, ALEA Lat. Am. J. Probab. Math. Stat. **21** (2024), no. 1, 73–94, http://arxiv.org/abs/2304.12975. MR4703770

[10] Amir Dembo and Ofer Zeitouni, *Large deviations techniques and applications*, second ed., Applications of Mathematics (New York), vol. 38, Springer-Verlag, New York, 1998. MR1619036

[11] B. Derrida, E. Domany, and D. Mukamel, *An exact solution of a one-dimensional asymmetric exclusion model with open boundaries*, J. Statist. Phys. **69** (1992), no. 3-4, 667–687. MR1193854

[12] B. Derrida, C. Enaud, and J. L. Lebowitz, *The asymmetric exclusion process and Brownian excursions*, J. Statist. Phys. **115** (2004), no. 1-2, 365–382. MR2070099

[13] B. Derrida and M. R. Evans, *Exact correlation functions in an asymmetric exclusion model with open boundaries*, J. Physique I **3** (1993), no. 2, 311–322. MR1215845

[14] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, *Exact solution of a* $1d$ *asymmetric exclusion model using a matrix formulation*, J. Phys. A **26** (1993), no. 7, 1493–1517. MR1219679

[15] B. Derrida, J. L. Lebowitz, and E. R. Speer, *Shock profiles for the asymmetric simple exclusion process in one dimension*, J. Statist. Phys. **89** (1997), no. 1-2, 135–167, Dedicated to Bernard Jancovici. MR1492490

[16] B. Derrida, J. L. Lebowitz, and E. R. Speer, *Exact large deviation functional of a stationary open driven diffusive system: the asymmetric exclusion process*, J. Statist. Phys. **110** (2003), no. 3-6, 775–810, Special issue in honor of Michael E. Fisher's 70th birthday (Piscataway, NJ, 2001). MR1964689

[17] B Derrida, JL Lebowitz, and ER Speer, *Exact free energy functional for a driven diffusive open stationary nonequilibrium system*, Physical Review Letters **89** (2002), no. 3, 030601.

[18] Bernard Derrida, *Matrix ansatz and large deviations of the density in exclusion processes*, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 367–382. MR2275686

[19] Bernard Derrida, *Non-equilibrium steady states: fluctuations and large deviations of the density and of the current*, J. Stat. Mech. Theory Exp. **2007** (2007), no. 7, P07023, 45 pp. MR2335699

[20] Bernard Derrida, Martin R. Evans, Vincent Hakim, and Vincent Pasquier, *Exact solution of a* $1D$ *asymmetric exclusion model using a matrix formulation*, J. Phys. A **26** (1993), no. 7, 1493–1517. MR1219679

[21] Enrica Duchi and Gilles Schaeffer, *A combinatorial approach to jumping particles*, J. Combin. Theory Ser. A **110** (2005), no. 1, 1–29. MR2128962

[22] Paul Dupuis and Richard S. Ellis, *A weak convergence approach to the theory of large deviations*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1997, A Wiley-Interscience Publication. MR1431744

[23] Thomas M. Liggett, *Ergodic theorems for the asymmetric simple exclusion process*, Trans. Amer. Math. Soc. **213** (1975), 237–261. MR410986

[24] Evita Nestoridi and Dominik Schmid, *Approximating the stationary distribution of the ASEP with open boundaries*, Comm. Math. Phys. **405** (2024), no. 8, Paper No. 176, 64 pp. MR4777077

[25] Gunter Schütz and Eytan Domany, *Phase transitions in an exactly soluble one-dimensional exclusion process*, Journal of Statistical Physics **72** (1993), no. 1-2, 277–296.

[26] Yizao Wang, Jacek Wesołowski, and Zongrui Yang, *Askey-Wilson signed measures and open ASEP in the shock region*, Int. Math. Res. Not. IMRN (2024), no. 15, 11104–11134. MR4782793

[27] Yizao Wang and Zongrui Yang, *From asymmetric simple exclusion processes with open boundaries to stationary measures of open KPZ fixed point: the shock region*, 2024, https://arxiv.org/abs/2406.09252.

Acknowledgments. This research benefited from discussions with Guillaume Barraquand, Ivan Corwin, and Yizao Wang.