

Large Deviations for Some Corner Growth Models with Inhomogeneity

E. Emrah¹ and C. Janjigian²

¹Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, USA.
E-mail: eemrah@andrew.cmu.edu

²Laboratoire de Probabilités et Modèles Aléatoires, Université Paris Diderot - Paris 7.
Paris, France. E-mail: cjanjigian@math.univ-paris-diderot.fr

Received June 7, 2016, revised December 19, 2016

Abstract. We study an inhomogeneous generalization of the classical corner growth in which the weights are exponentially distributed with random parameters. Our interest is in the large deviation properties of the last passage times. We obtain tractable variational representations of the right tail large deviation rate functions in both the quenched and annealed settings and estimates for left tail large deviations. We also compute expansions of the right tail rate functions near the shape function, which are consistent with the expectation of KPZ type fluctuations in an appropriate regime.

KEYWORDS: corner growth model; directed last-passage percolation; TASEP; exactly solvable models; large deviations; rate functions

AMS SUBJECT CLASSIFICATION: 60K35; 60K37

1. Introduction

Let $W = \{W(i, j) : i, j \in \mathbb{N}\}$ be a collection of nonnegative real random variables (weights) with joint distribution P . We consider a growing random subset of \mathbb{N}^2 . Initially, this subset is empty and the growth rule is as follows: at time zero, we start a countdown of length $W(1, 1)$; when this countdown ends, we add $(1, 1)$ to the subset. This process then iterates: once the bottom and left neighbors (if they exist) of a site (i, j) have been added, a countdown of length $W(i, j)$ begins; when it ends, (i, j) joins the subset.

Our interest is in the last-passage times $G = \{G(i, j) : i, j \in \mathbb{N}\}$, defined recursively by

$$G(i, j) = G(i - 1, j) \vee G(i, j - 1) + W(i, j), \quad G(i, 0) = G(0, j) = 0, \quad (1.1)$$

for $i, j \in \mathbb{N}$. These random variables encode the evolution of the set described above in the sense that $(i, j) \in \mathbb{N}^2$ is added to the set at time $t = G(i, j)$. They can also be viewed as describing a directed last-passage percolation model since (1.1) is equivalent to

$$G(m, n) = \max_{\pi \in \Pi_{(1,1),(m,n)}} \sum_{(i,j) \in \pi} W(i, j), \tag{1.2}$$

where $\Pi_{(k,l),(m,n)}$ is the set of all directed paths from (k, l) to (m, n) . A directed path $\pi \in \Pi_{(k,l),(m,n)}$ is a finite sequence $(u_i, v_i)_{i \in [p]}$ in \mathbb{Z}^2 such that $(u_1, v_1) = (k, l)$, $(u_p, v_p) = (m, n)$ and $(u_{i+1} - u_i, v_{i+1} - v_i) \in \{(1, 0), (0, 1)\}$ for $1 \leq i < p$.

The corner growth model also maps to a generalization of the totally asymmetric exclusion process (TASEP) on \mathbb{Z} started from step initial conditions. In this interpretation, we begin with particles at the sites $i < 0$ and holes at the sites $i \geq 0$. We label the particles with $i \in \mathbb{N}$, counting from right to left and holes with $j \in \mathbb{N}$ from left to right. In the dynamics, particle i and hole j interchange at time $G(i, j)$. If we denote the position of particle i at time t by $\sigma(i, t)$ then we have

$$\sigma(i, t) = -i + \max\{j \in \mathbb{N} : G(i, j) \leq t\}. \tag{1.3}$$

If the weights are i.i.d. with geometric or exponential marginals, this process is the usual TASEP run in discrete or continuous time, respectively [25, p. 5].

These and related models have received substantial research attention in the past two decades, partially in connection with KPZ universality. See the surveys [5, 18]. When P is i.i.d. with geometric or exponential marginals, it has been possible to compute various statistics of the last-passage times. For example

$$\lim_{n \rightarrow \infty} \frac{G(\lfloor ns \rfloor, \lfloor nt \rfloor)}{n} = m(s + t) + 2\sigma\sqrt{st} \quad \text{for } s, t > 0 \text{ } P\text{-a.s.}, \tag{1.4}$$

where m and σ^2 are the common mean and the variance of the weights. The exponential case of (1.4) was first proved in [21] and the geometric case appeared in [3, 15, 23]. Large deviation principles for the last-passage times were derived in [14, 22]. These papers identified the right-tail rate function and the correct decay rate for both the right and left tails. It is also established in [14] that the model exhibits KPZ statistics; the fluctuations around the limit in (1.4) are of order $n^{1/3}$ and appropriately rescaled last-passage times converge to the Tracy–Widom GUE distribution.

This paper concerns an inhomogeneous generalization of the classical i.i.d. exponential model. Given parameter sequences $\mathbf{a} = (a_n)_{n \in \mathbb{N}}$ and $\mathbf{b} = (b_n)_{n \in \mathbb{N}}$ taking values in $(0, \infty)$, we define a measure under which the weights are independent and $W(i, j)$ is exponentially distributed with mean $(a_i + b_j)^{-1}$. We

state the assumptions of our model precisely in Section 2, but for the moment we will outline our results in the case that \mathbf{a} and \mathbf{b} are independent i.i.d. sequences which are bounded away from zero and have finite means. We refer to the model where we condition on these random sequences as the quenched model. In the framework of particle systems, the quenched model corresponds to an inhomogeneous continuous time TASEP with particlewise and holewise inhomogeneity. The annealed measure is constructed by averaging the family of quenched measures over the joint distribution of (\mathbf{a}, \mathbf{b}) . Weights in the quenched model are independent, but not identically distributed in general; in the annealed model, the weights are identically distributed but correlated along rows and columns.

For the models we consider, the almost sure limit of $n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor)$ is deterministic and can be characterized as the solution of a certain variational problem, see [8]. With some choices of the marginal distributions of \mathbf{a} and \mathbf{b} , this family of models has a feature not present in the previous exactly solvable corner growth models: the presence of linear segments of the shape function. This is of some interest because different fluctuation exponents are expected in these different regions if the weights are weakly correlated. For a particle system perspective on this phenomenon in the case where \mathbf{a} is almost surely constant and \mathbf{b} is i.i.d. and bounded, see [27].

The present paper is devoted to the question of large deviations corresponding (1.4). In the quenched setting, we are able to prove a large deviation principle with rate n and a rate function given by the solution to a reasonably tractable variational problem. With certain choices of the weights and in certain directions, we provide some explicit formulas for these rate functions.

In the annealed setting, we have a variational expression for the right tail rate function which is similar to the variational expression in the quenched setting, though we no longer have any non-trivial explicitly computable examples. Deviations to the right of the shape function in the annealed model are connected to deviations in the quenched model through a variational problem involving relative entropy. Heuristically, these deviations should arise from perturbations of (\mathbf{a}, \mathbf{b}) combined with deviations in the quenched model with these perturbed parameters. There are rate n annealed large deviations to the left of the shape function. This is in contrast to the i.i.d. models, where the rate is n^2 [14, 22]. We show that this occurs by using the fact that it is possible to see a finite entropy deviation of the (order n many) parameters $\{a_i\}_{i=1}^{\lfloor ns \rfloor}$ and $\{b_j\}_{j=1}^{\lfloor nt \rfloor}$ which affect the distribution of $G(\lfloor ns \rfloor, \lfloor nt \rfloor)$ and makes the shape function smaller.

We identify the expansions of both the quenched and annealed rate functions near the shape function. In the quenched model, for directions in which the shape function is strictly concave, these expansions are heuristically consistent with the expectation of Tracy–Widom GUE fluctuations. Fluctuation results for an inhomogeneous version of the closely related Seppäläinen–Johansson model (oriented digital boiling) were previously obtained in a series of papers by Gravner, Tracy, and Widom [10–12]. The question of what happens at the

interface of the linear and concave regions was left open in those papers. At the interface of the linear and concave regions for the models we study, our results suggest KPZ type fluctuations under a moment condition. We elaborate on this connection in the next section.

To prove the variational formulas for the right tail rate functions, we follow an approach introduced in [24] and applied in [9, 13, 22]. The key technical condition making this scheme tractable is an analogue of Burke's theorem from queueing theory, which in this setting corresponds to the existence of a stationary version of the model, as discussed in Proposition 2.1. We expect that the techniques employed in this paper could be used to obtain similar results in inhomogeneous versions of other models with the Burke property, such as the log gamma polymer [26], the strict-weak polymer [6] and the corner growth model with geometric weights [8].

The principal contributions of this paper are as follows. To the best of our knowledge, this is the first inhomogeneous model in the KPZ class for which exact large deviation rate functions have been computed. The rate functions we obtain in both the quenched and annealed settings are tractable. For general choices of the distributions of the sequences (\mathbf{a}, \mathbf{b}) , we identify the asymptotic rate that the right tail rate function tends to zero near the shape function, suggesting KPZ type fluctuations for the quenched model in appropriate directions. In particular, our results suggest a partial answer to the problem of what type of fluctuations to expect at the interface of the linear and concave regions. We further connect our quenched and annealed rate functions through a natural variational problem involving relative entropy.

The paper is organized as follows. In Section 2, we define the model precisely and state our results. The remaining sections are devoted to proofs. In Section 3, we discuss the stationary model and compute the Lyapunov exponents of the last passage times. In Section 4, we study the extremizers of the variational problems for the rate functions and Lyapunov exponents. We then estimate the probability of left tail large deviations in Section 5. In Section 6 we show that the Legendre–Fenchel transform of the right tail rate function is given by the previously computed Lyapunov exponents. These results are combined to prove the large deviation principle for the quenched model. In Section 7, we note that the extremizers for the annealed model are connected to the extremizers for a quenched model with different parameters, which gives a variational connection between the quenched and annealed rate functions. Understanding of the extremizers also allows us to prove the scaling estimates in Section 8. We include the standard subadditivity arguments showing existence and regularity of the Lyapunov exponents and right tail rate functions in Appendix A.

Notation. For real numbers a, b , we denote $\max(a, b) = a \vee b$ and $\min(a, b) = a \wedge b$. We take the convention that $\mathbb{N} = \{n \in \mathbb{Z} : n > 0\}$ and $\mathbb{R}_+ = \{x \in \mathbb{R} : x > 0\}$. For $D \subset \mathbb{R}$, we denote by $\mathcal{M}_1(D)$ the collection of probability measures on D . For $\eta \in \mathcal{M}_1(D)$, we use the notation $\underline{\eta} = \text{ess-inf}\{\eta\}$ and $\bar{\eta} = \text{ess-sup}\{\eta\}$.

Given $\nu, \mu \in \mathcal{M}_1(D)$, the relative entropy is defined by

$$H(\nu|\mu) = E^\nu \log(d\nu/d\mu)$$

if ν is absolutely continuous with respect to μ and ∞ otherwise. See for example the discussions in [7] and [19] for basic properties of the relative entropy. We denote absolute continuity of ν with respect to μ by $\nu \ll \mu$. For probability measures ν, μ with $\nu \ll \mu$, we write $(d\nu/d\mu)(x) \simeq f(x)$ if $\nu(dx) = [\int f(x)\mu(dx)]^{-1}f(x)\mu(dx)$. For a probability measure μ on \mathbb{R}_+ , define

$$\mathcal{M}^\mu = \{\nu \in \mathcal{M}_1(\mathbb{R}_+) : H(\nu|\mu) < \infty\} \tag{1.5}$$

and note that for each μ , \mathcal{M}^μ is a convex set by convexity of $H(\nu|\mu)$.

For $f : \mathbb{R} \rightarrow (-\infty, \infty]$, $f^*(\xi) = \sup_{x \in \mathbb{R}} \{x\xi - f(x)\}$ defines the Legendre–Fenchel transform. We refer the reader to [20] for basic properties of this transform.

2. Model and results

2.1. Model

Denote by $W(i, j)$ the projection $\mathbb{R}_+^{\mathbb{N}^2} \rightarrow \mathbb{R}_+$ onto the coordinate (i, j) for $i, j \in \mathbb{N}$. For any sequences $\mathbf{a} = (a_1, a_2, \dots), \mathbf{b} = (b_1, b_2, \dots)$ taking values in \mathbb{R}_+ , we define $\mathbf{P}_{\mathbf{a}, \mathbf{b}}$ to be the product measure on $\mathbb{R}_+^{\mathbb{N}^2}$ satisfying

$$\mathbf{P}_{\mathbf{a}, \mathbf{b}}(W(i, j) \geq x) = e^{-(a_i + b_j)x} \quad \text{for } i, j \in \mathbb{N} \text{ and } x \geq 0.$$

We will draw the sequences (\mathbf{a}, \mathbf{b}) randomly from a distribution μ on $\mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}}$. For $k \in \mathbb{Z}_+$, let τ_k denote the shift $(c_n)_{n \in \mathbb{N}} \mapsto (c_{n+k})_{n \in \mathbb{N}}$. In all of the results that follow, we make the following assumptions on (\mathbf{a}, \mathbf{b}) . We assume that \mathbf{a} and \mathbf{b} are stationary sequences under μ . We assume further that μ is separately ergodic with respect to $\tau_k \times \tau_l$ for $k, l \in \mathbb{N}$. This means that if $k, l \in \mathbb{N}$ and $B \subset \mathbb{R}_+^{\mathbb{N}} \times \mathbb{R}_+^{\mathbb{N}}$ is a Borel set with $(\tau_k \times \tau_l)^{-1}(B) = B$ then $\mu(B) \in \{0, 1\}$.

The annealed distribution \mathbb{P} is given by $\mathbb{P}(B) = E[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(B)]$ for any Borel set $B \subset \mathbb{R}_+^{\mathbb{N}^2}$, where E is the expectation under μ . Let $\mathbf{E}_{\mathbf{a}, \mathbf{b}}$ and \mathbb{E} denote the expectations under $\mathbf{P}_{\mathbf{a}, \mathbf{b}}$ and \mathbb{P} , respectively. We denote by α and β the distributions of a_1 and b_1 and take the convention that a and b are random variables with distributions α and β respectively. In all of the following results, we will assume that $E[a + b] < \infty$ and $\underline{\alpha} + \underline{\beta} > 0$. Finally, all large deviation results under \mathbb{P} are limited to the case in which \mathbf{a} and \mathbf{b} are independent i.i.d. sequences.

We will also consider a ‘stationary’ model defined on the extended sample space $\mathbb{R}_+^{\mathbb{Z}_+^2}$. Each weight $W(i, j)$ is now redefined as the projection onto coordi-

nate (i, j) for $i, j \in \mathbb{Z}_+^2$. Introduce the last-passage times

$$\hat{G}(m, n) = \max_{\pi \in \Pi_{(0,0),(m,n)}} \sum_{i,j \in \pi} W(i, j) \quad \text{for } m, n \in \mathbb{Z}_+. \tag{2.1}$$

For sequences \mathbf{a} and \mathbf{b} in $(0, \infty)$ and $z \in (-\underline{\alpha}, \underline{\beta})$, define the product measure $\mathbf{P}_{\mathbf{a},\mathbf{b}}^z$ on $\mathbb{R}_+^{\mathbb{Z}_+^2}$ by

$$\begin{aligned} \mathbf{P}_{\mathbf{a},\mathbf{b}}^z(W(i, j) \geq x) &= \exp(-(a_i + b_j)x), & \mathbf{P}_{\mathbf{a},\mathbf{b}}^z(W(0, 0) = 0) &= 1, \\ \mathbf{P}_{\mathbf{a},\mathbf{b}}^z(W(i, 0) \geq x) &= \exp(-(a_i + z)x), & \mathbf{P}_{\mathbf{a},\mathbf{b}}^z(W(0, j) \geq x) &= \exp(-(b_j - z)x) \end{aligned} \tag{2.2}$$

for $i, j \in \mathbb{Z}_+$ and $x \geq 0$. We will use definition (2.2) for $z = -\underline{\alpha}$ when $a_i > \underline{\alpha}$ for $i \in \mathbb{N}$ and for $z = \underline{\beta}$ when $b_j > \underline{\beta}$ for $j \in \mathbb{N}$. The utility of these measures is that the last-passage increments given by $I(m, n) = \hat{G}(m, n) - \hat{G}(m - 1, n)$ for $m \geq 1, n \geq 0$ and $J(m, n) = \hat{G}(m, n) - \hat{G}(m, n - 1)$ for $m \geq 0, n \geq 1$ are stationary in the following sense.

Proposition 2.1 (Proposition 4.1 in [8]). *Let $k, l \in \mathbb{Z}_+$. Under $\mathbf{P}_{\mathbf{a},\mathbf{b}}^z$,*

- (a) $I(i, l)$ has the same distribution as $W(i, 0)$ for $i \in \mathbb{N}$.
- (b) $J(k, j)$ has the same distribution as $W(0, j)$ for $j \in \mathbb{N}$.
- (c) The random variables $\{I(i, l) : i > k\} \cup \{J(k, j) : j > l\}$ are jointly independent.

For admissible z , define the measure \mathbb{P}^z on $\mathbb{R}_+^{\mathbb{Z}_+^2}$ by $\mathbb{P}^z(B) = \mathbf{E}[\mathbf{P}_{\mathbf{a},\mathbf{b}}^z(B)]$ for any Borel set B . Let $\mathbf{E}_{\mathbf{a},\mathbf{b}}^z$ and \mathbb{E}^z denote the expectations under $\mathbf{P}_{\mathbf{a},\mathbf{b}}^z$ and \mathbb{P}^z , respectively.

2.2. Results

We begin by briefly summarizing the results from [8]. The ergodicity assumptions on μ and the superadditivity of the last-passage times imply that $\lim_{n \rightarrow \infty} n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) = g(s, t)$ for $s, t > 0$ \mathbb{P} -a.s. and $\mathbf{P}_{\mathbf{a},\mathbf{b}}$ -a.s. for μ -a.e. (\mathbf{a}, \mathbf{b}) for some deterministic function g known as the *shape function*. g admits the variational representation

$$g(s, t) = \inf_{z \in [-\underline{\alpha}, \underline{\beta}]} \left\{ s \mathbf{E} \left[\frac{1}{a + z} \right] + t \mathbf{E} \left[\frac{1}{b - z} \right] \right\} \quad \text{for } s, t > 0. \tag{2.3}$$

The infimum above is actually a minimum with a unique minimizer and the function g_z given by $g_z(s, t) = s \mathbf{E} [(a + z)^{-1}] + t \mathbf{E} [(b - z)^{-1}]$ is the shape function in the stationary version of the model. At times we will also view $g(s, t)$ as a function of $(\alpha, \beta) \in \mathcal{M}_1(\mathbb{R}_+)^2$. In these cases, we will use the notation

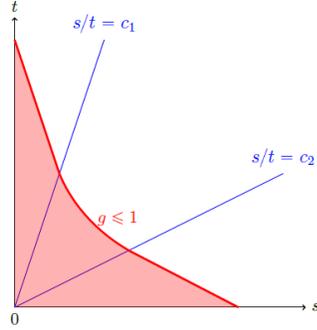


Figure 1: An illustration of the sublevel set $g \leq 1$ and the rays $s/t = c_1$ and $s/t = c_2$ when $0 < c_1 < c_2 < \infty$.

$(\alpha, \beta) \mapsto g(s, t) \equiv g_{\alpha, \beta}(s, t)$ to highlight the dependence on these measures. This map will be considered for any $(\alpha, \beta) \in \mathcal{M}_1(\mathbb{R}_+)^2$.

Set

$$c_1 = \frac{\mathbb{E}[(b + \underline{\alpha})^{-2}]}{\mathbb{E}[(a - \underline{\alpha})^{-2}]} \quad c_2 = \frac{\mathbb{E}[(b - \underline{\beta})^{-2}]}{\mathbb{E}[(a + \underline{\beta})^{-2}]} \tag{2.4}$$

Then $0 \leq c_1 < c_2 \leq \infty$, and $c_1 = 0$ and $c_2 = \infty$ if and only if $\mathbb{E}[(a - \underline{\alpha})^{-2}] = \infty$ and $\mathbb{E}[(b - \underline{\beta})^{-2}] = \infty$, respectively. It can be seen from (2.3) that g is strictly concave for $c_1 < s/t < c_2$ and is linear for $s/t \leq c_1$ or $s/t \geq c_2$, see Figure 1.

We show in Proposition A.2 of the appendix that for $s, t, \lambda > 0$, we may define the quenched and annealed Lyapunov exponents by

$$\begin{aligned} \mathbf{L}_{s,t}(\lambda) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] \quad \mu\text{-a.s.}, \\ \mathbb{L}_{s,t}(\lambda) &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}]. \end{aligned}$$

Our first result is an exact computation of these exponents.

Theorem 2.1. For $s, t, \lambda > 0$, if $0 < \lambda \leq \underline{\alpha} + \underline{\beta}$,

$$\mathbf{L}_{s,t}(\lambda) = \inf_{z \in [-\underline{\alpha}, \underline{\beta} - \lambda]} \left\{ s \mathbb{E} \log \left(\frac{a + z + \lambda}{a + z} \right) + t \mathbb{E} \log \left(\frac{b - z}{b - z - \lambda} \right) \right\}; \tag{2.5}$$

if $\lambda > \underline{\alpha} + \underline{\beta}$, $\mathbf{L}_{s,t}(\lambda) = \infty$. If $0 < \lambda \leq \underline{\alpha} + \underline{\beta}$,

$$\mathbb{L}_{s,t}(\lambda) = \inf_{z \in [-\underline{\alpha}, \underline{\beta} - \lambda]} \left\{ s \log \mathbb{E} \left(\frac{a + z + \lambda}{a + z} \right) + t \log \mathbb{E} \left(\frac{b - z}{b - z - \lambda} \right) \right\}; \tag{2.6}$$

if $\lambda > \underline{\alpha} + \underline{\beta}$, $\mathbb{L}_{s,t}(\lambda) = \infty$.

Having proven Theorem 2.1, a proof similar to the proof of Theorem 2.1 allows us to compute the Lyapunov exponents in the stationary version of the model.

Theorem 2.2. *For $z \in (-\underline{\alpha}, \underline{\beta})$, almost surely for all $s, t > 0$ and $\lambda \in (0, (\underline{\alpha} + z) \wedge (\underline{\beta} - z))$*

$$\begin{aligned} \mathbf{L}_{s,t}^z(\lambda) &:= \lim_{n \rightarrow \infty} n^{-1} \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z [e^{\lambda \hat{G}(\lfloor ns \rfloor, \lfloor nt \rfloor)}] \\ &= \left\{ s \mathbf{E} \left[\log \frac{a+z}{a+z-\lambda} \right] + t \mathbf{E} \left[\log \frac{b-z+\lambda}{b-z} \right] \right\} \\ &\quad \vee \left\{ s \mathbf{E} \left[\log \frac{a+z+\lambda}{a+z} \right] + t \mathbf{E} \left[\log \frac{b-z}{b-z-\lambda} \right] \right\}. \end{aligned}$$

Similarly, we show in Proposition A.1 that for $s, t > 0$ and $r \in \mathbb{R}$, we may define right tail rate functions by

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) &= \mathbf{J}_{s,t}(r) \quad \mu\text{-a.s.}, \\ \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) &= \mathbb{J}_{s,t}(r) \end{aligned}$$

Using Theorem 2.1, we show that

Theorem 2.3. *For $s, t > 0$,*

$$\mathbf{J}_{s,t}(r) = \sup_{\substack{\lambda \in (0, \underline{\alpha} + \underline{\beta}] \\ z \in [-\underline{\alpha}, \underline{\beta} - \lambda]}} \left\{ r\lambda - s \mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right) - t \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \right\} \quad (2.7)$$

for $r \geq g(s, t)$, and $\mathbf{J}_{s,t}(r) = 0$ for $r < g(s, t)$.

$$\mathbb{J}_{s,t}(r) = \sup_{\substack{\lambda \in (0, \underline{\alpha} + \underline{\beta}] \\ z \in [-\underline{\alpha}, \underline{\beta} - \lambda]}} \left\{ r\lambda - s \log \mathbf{E} \left[\frac{a+z+\lambda}{a+z} \right] - t \log \mathbf{E} \left[\frac{b-z}{b-z-\lambda} \right] \right\} \quad (2.8)$$

for $r \geq g(s, t)$, and $\mathbb{J}_{s,t}(r) = 0$ for $r < g(s, t)$.

The preceding result also describes left tail large deviations for a tagged particle in an inhomogeneous TASEP with step initial condition. This TASEP can be obtained from the corner growth by defining the position of particle $i \in \mathbb{N}$ at time $t \geq 0$ as in (1.3). By monotonicity of G , the particles remain ordered i.e. $\sigma(i, t) > \sigma(i + 1, t)$ for $i \in \mathbb{N}$ and $t \geq 0$. Initially, $\sigma(i, 0) = -i$ for $i \in \mathbb{N}$ and particles move on \mathbb{Z} over time according to the following rule. If particle i is at site $-i + j - 1$, as soon as site $-i + j$ is vacant, particle i moves to site $-i + j$ after $W(i, j)$ amount of time. Since $\{\sigma(i, t) > j\} = \{G(i, i + j) < t\}$ as events, Theorem 2.3 implies the next corollary.

Corollary 2.1. For $x, y, t > 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{P}_{\mathbf{a}, \mathbf{b}}(\sigma(\lfloor nx \rfloor, nt) > \lfloor ny \rfloor) &= \mathbf{J}_{x, x+y}(t) \quad \text{a.s.}, \\ \lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(\sigma(\lfloor nx \rfloor, nt) > \lfloor ny \rfloor) &= \mathbb{J}_{x, x+y}(t). \end{aligned}$$

As with the shape function, we will at times consider the maps $(\alpha, \beta) \mapsto \mathbf{J}_{s,t}(r) \equiv \mathbf{J}_{s,t}^{\alpha, \beta}(r)$ and $(\alpha, \beta) \mapsto \mathbb{J}_{s,t}(r) \equiv \mathbb{J}_{s,t}^{\alpha, \beta}(r)$.

Note that the Lyapunov exponents and the right tail rate functions depend on μ only through the marginal distributions α and β . The variational problem in (2.7) can be solved exactly for certain choices of α, β, s and t . We note that if $r \geq g(s, t)$ and there exists $\lambda_\star \in (0, \underline{\alpha} + \underline{\beta})$ and $z_\star \in (-\underline{\alpha}, \underline{\beta} - \lambda_\star)$ such that

$$\begin{aligned} 0 &= s \mathbb{E} \left[\frac{1}{a + z_\star + \lambda_\star} - \frac{1}{a + z_\star} \right] + t \mathbb{E} \left[\frac{1}{b - z_\star - \lambda_\star} - \frac{1}{b - z_\star} \right], \\ r &= s \mathbb{E} \left[\frac{1}{a + z_\star + \lambda_\star} \right] + t \mathbb{E} \left[\frac{1}{b - z_\star - \lambda_\star} \right], \end{aligned}$$

then

$$\mathbf{J}_{s,t}(r) = \lambda_\star r - s \mathbb{E} \log \left(\frac{a + z_\star + \lambda_\star}{a + z_\star} \right) + t \mathbb{E} \log \left(\frac{b - z_\star}{b - z_\star - \lambda_\star} \right). \quad (2.9)$$

Example 2.1. If $\alpha = \beta = \delta_{c/2}$ for $c > 0$, then for $r \geq g(s, t) = c^{-1}(\sqrt{s} + \sqrt{t})^2$,

$$\mathbf{J}_{s,t}(r) = \sqrt{(s + t - cr)^2 - 4st} - 2s \cosh^{-1} \left(\frac{s - t + cr}{2\sqrt{csr}} \right) - 2t \cosh^{-1} \left(\frac{t - s + cr}{2\sqrt{ctr}} \right), \quad (2.10)$$

which recovers [22, Theorem 4.4].

Example 2.2. If $\alpha = \beta = p\delta_c + q\delta_d$ for $p, q, c, d > 0$ with $p + q = 1$ and $s = t$, then for $r \geq g(s, s) = 2s(p c^{-1} + q d^{-1})$,

$$\begin{aligned} \mathbf{J}_{s,s}(r) &= r \lambda_\star - sp \log \left(\frac{c + z_\star + \lambda_\star}{c + z_\star} \right) - tq \log \left(\frac{c - z_\star}{c - z_\star - \lambda_\star} \right) \\ &\quad - sq \log \left(\frac{d + z_\star + \lambda_\star}{d + z_\star} \right) - tq \log \left(\frac{d - z_\star}{d - z_\star - \lambda_\star} \right) \end{aligned}$$

where

$$\begin{aligned} z_\star &= \frac{2cp + 2dq + c^2r + d^2r - \sqrt{\Delta}}{2r}, & z_\star + \lambda_\star &= \frac{2cp + 2dq + c^2r + d^2r + \sqrt{\Delta}}{2r}, \\ \Delta &= (2cp + 2dq + c^2r + d^2r)^2 + 4r(2cd^2p + 2c^2dq - c^2d^2r). \end{aligned}$$

More complicated exact formulas in this model are available in all directions (s, t) .

Example 2.3. If α and β are uniform on $[c/2, c/2 + l]$ for $c, l > 0$ and $s = t$, then

$$\mathbf{J}_{s,s}(r) = r \lambda_\star - \frac{2s}{l} \int_{c/2}^{c/2+l} \log\left(\frac{x + z_\star + \lambda_\star}{x + z_\star}\right) dx \text{ for } r \geq g(s, s) = \frac{2s}{l} \log\left(1 + \frac{2l}{c}\right),$$

where

$$z_\star = -\sqrt{\frac{(c/2 + l)^2 - c^2 e^{rl/s}/4}{1 - e^{rl/s}}}, \quad z_\star + \lambda_\star = \sqrt{\frac{(c/2 + l)^2 - c^2 e^{rl/s}/4}{1 - e^{rl/s}}}.$$

Left tail large deviations in the quenched model have rate strictly larger than n . We expect that under mild hypotheses the correct rate should be n^2 , as is the case in the homogeneous model where $\alpha = \beta = \delta_{c/2}$ [14, 22].

Lemma 2.1. $\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{P}_{\mathbf{a}, \mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \leq nr) = \infty$ for $s, t > 0$ and $r < g(s, t)$ μ -a.s.

Combining our results for the right and left tail deviations, we can prove a full quenched LDP at rate n . The rate function is given by

$$\mathbf{I}_{s,t}(r) = \begin{cases} \mathbf{J}_{s,t}(r) & r \geq g(s, t), \\ \infty & r < g(s, t). \end{cases} \tag{2.11}$$

As before, we will at times use the notation $(\alpha, \beta) \mapsto \mathbf{I}_{s,t}(r) \equiv \mathbf{I}_{s,t}^{\alpha,\beta}(r)$.

Theorem 2.4. μ -a.s, for any $s, t > 0$, the distribution of $n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor)$ under $\mathbf{P}_{\mathbf{a}, \mathbf{b}}$ satisfies a large deviation principle with rate n and convex, good rate function $\mathbf{I}_{s,t}$.

Although our proof of the large deviation principle goes through the Lyapunov exponents, we do not apply the Gärtner – Ellis theorem. The steepness condition in this model is $E[(a - \alpha)^{-1}] = E[(b - \beta)^{-1}] = \infty$, which would rule out having linear segments of the shape function and so is too restrictive.

In contrast to the quenched case, there are non-trivial annealed large deviations at rate n . The following bound gives a mechanism for these deviations.

Lemma 2.2. For any $x < y$,

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \in (x, y)) \\ \leq \inf_{\substack{\nu_1 \in \mathcal{M}^\alpha, \nu_2 \in \mathcal{M}^\beta \\ g_{\nu_1, \nu_2}(s, t) \in (x, y)}} \{s H(\nu_1 | \alpha) + t H(\nu_2 | \beta)\}. \end{aligned}$$

The other bound needed to show that n is the correct rate for certain left tail large deviations follows from essentially the same argument used to show that the quenched rate is strictly larger than n . This is discussed briefly after the proof of Lemma 2.1. To show that there are rate n annealed left tail large deviations it suffices to show that there exist $\nu_1 \in \mathcal{M}^\alpha$ and $\nu_2 \in \mathcal{M}^\beta$ with $g_{\nu_1, \nu_2}(s, t) < g_{\alpha, \beta}(s, t)$. We give a simple proof that under mild conditions this is the case in Lemma 5.1. We expect that this mechanism is not sharp.

Example 2.4. Suppose that $\alpha = \delta_1/2 + \delta_2/2$ and $\beta = \delta_1$, and recall that $\mathcal{M}^\alpha = \{p\delta_1 + (1-p)\delta_2 : 0 \leq p \leq 1\}$. For $0 \leq p \leq 1$, call $\alpha_p = p\delta_1 + (1-p)\delta_2$. Then $\{g_{\alpha_p, \beta}(1, 9) : 0 \leq p \leq 1\} = \{5.3\} \cup (5.5, 8]$. The reason for the discontinuity in this example is that if $p > 0$, then the functional in (2.3) is minimized on the set $(-1, 1)$, but if $p = 0$, the minimization occurs on $(-2, 1)$. We have chosen $s = 1, t = 9$ so that the minimizer for the $p = 0$ case occurs in $(-2, -1)$. The bound one obtains from Lemma 2.2 in this example is infinite when applied to the interval $(5.4, 5.5)$. The finite relative entropy perturbation of the a_i parameters switching the distribution to δ_2 turns this into a right tail large deviation.

The next theorem connects quenched rate function and annealed right tail rate function through a variational problem. We expect that this result means that large deviations above the shape function in the annealed model with marginals α and β can be viewed as a large deviation in the parameters $\{a_i\}_{i=1}^{\lfloor ns \rfloor}$ and $\{b_j\}_{j=1}^{\lfloor nt \rfloor}$ which affect the distribution of $G(\lfloor ns \rfloor, \lfloor nt \rfloor)$, followed by a deviation in the quenched model with these perturbed parameters. Our proof is purely analytic and does not show this interpretation directly. A similar, but stronger, connection was shown for random walk in a random environment by Comets, Gantert and Zeitouni in [4].

Theorem 2.5. For any $s, t > 0$ and $r > g(s, t)$,

$$\mathbb{J}_{s,t}^{\alpha,\beta}(r) = \inf_{\substack{\nu_1 \in \mathcal{M}^\alpha \\ \nu_2 \in \mathcal{M}^\beta}} \{ \mathbf{I}_{s,t}^{\nu_1, \nu_2}(r) + s \mathbf{H}(\nu_1 | \alpha) + t \mathbf{H}(\nu_2 | \beta) \}.$$

A minimizing pair (ν_1, ν_2) exists and the equality

$$\mathbb{J}_{s,t}^{\alpha,\beta}(r) = \mathbf{I}_{s,t}^{\nu_1, \nu_2}(r) + s \mathbf{H}(\nu_1 | \alpha) + t \mathbf{H}(\nu_2 | \beta)$$

holds if and only if

$$\frac{d\nu_1}{d\alpha}(a) \simeq \frac{a + z_\star + \lambda_\star}{a + z_\star}, \quad \frac{d\nu_2}{d\beta}(b) \simeq \frac{b - z_\star}{b - z_\star - \lambda_\star}$$

where z_\star and λ_\star are the unique z_\star, λ_\star with $\lambda_\star \in [0, \alpha + \beta]$, $z_\star \in [-\alpha, \beta - \lambda_\star]$ satisfying

$$\mathbb{J}_{s,t}^{\alpha,\beta}(r) = r \lambda_\star - s \log \mathbf{E}^\alpha \left[\frac{a + z_\star + \lambda_\star}{a + z_\star} \right] - t \log \mathbf{E}^\beta \left[\frac{b - z_\star}{b - z_\star - \lambda_\star} \right]. \tag{2.12}$$

It is natural to conjecture that this variational connection describes all rate n annealed large deviations, rather than just annealed right tail large deviations. We have been unable to prove this result.

The next result concerns the regularity of our rate functions. Our rate functions are convex and differentiable to the right of $g(s, t)$, but we note that for certain choices of α and β they can have linear segments; see Lemma 4.4 and the comments preceding it.

Theorem 2.6. *For any $s, t > 0$, both $\mathbf{J}_{s,t}$ and $\mathbb{J}_{s,t}$ are continuously differentiable on $[g(s, t), +\infty)$.*

Finally, we describe the leading order asymptotics of $\mathbf{J}_{s,t}(r)$ and $\mathbb{J}_{s,t}(r)$ as $r \downarrow g(s, t)$ and comment on the implications for the fluctuations of the last-passage times. Let ζ denote the unique minimizer of (2.3).

Theorem 2.7. *For any $s, t > 0$, as $\varepsilon \downarrow 0$,*

$$\mathbf{J}_{s,t}(g(s, t) + \varepsilon) = \begin{cases} \left(-s \mathbb{E} \left[\frac{2}{(a - \alpha)^2} \right] + t \mathbb{E} \left[\frac{2}{(b + \alpha)^2} \right] \right)^{-1} \varepsilon^2 + o(\varepsilon^2) \\ \text{if } s/t < c_1, \\ \frac{2}{3} \left(s \mathbb{E} \left[\frac{1}{(a - \alpha)^3} \right] + t \mathbb{E} \left[\frac{1}{(b + \alpha)^3} \right] \right)^{-1/2} \varepsilon^{3/2} + o(\varepsilon^{3/2}) \\ \text{if } s/t = c_1 \text{ and } \mathbb{E}[(a - \alpha)^{-3}] < \infty, \\ \frac{4}{3} \left(s \mathbb{E} \left[\frac{1}{(a + \zeta)^3} \right] + t \mathbb{E} \left[\frac{1}{(b - \zeta)^3} \right] \right)^{-1/2} \varepsilon^{3/2} + o(\varepsilon^{3/2}) \\ \text{if } c_1 < s/t < c_2, \\ [4pt] \frac{2}{3} \left(s \mathbb{E} \left[\frac{1}{(a + \beta)^3} \right] + t \mathbb{E} \left[\frac{1}{(b - \beta)^3} \right] \right)^{-1/2} \varepsilon^{3/2} + o(\varepsilon^{3/2}) \\ \text{if } s/t = c_2 \text{ and } \mathbb{E}[(b - \beta)^{-3}] < \infty, \\ \left(s \mathbb{E} \left[\frac{2}{(a + \beta)^2} \right] - t \mathbb{E} \left[\frac{2}{(b - \beta)^2} \right] \right)^{-1} \varepsilon^2 + o(\varepsilon^2) \\ \text{if } s/t > c_2. \end{cases}$$

We do not have an intuitive explanation for the presence of an extra factor of $1/2$ in the boundary cases $s/t = c_1, c_2$.

The results of Theorem 2.7 in the concave region S and the boundary lines $s/t = c_1$ or c_2 are heuristically consistent with the expectation of KPZ type fluctuations. For example, to see this set

$$C = s \mathbb{E} \left[\frac{1}{(a + \zeta)^3} \right] + t \mathbb{E} \left[\frac{1}{(b - \zeta)^3} \right] = \frac{1}{2} \partial_z^2 g_z(s, t) \Big|_{z=\zeta}$$

and assume that our asymptotic result in the concave region hold for finite n . Then for $(s, t) \in S$ and large r , we expect to see

$$\begin{aligned} \mathbf{P}_{\mathbf{a}, \mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) - ng(s, t) \geq n^{1/3} C^{1/3} r) &\approx \exp\left\{-\frac{4}{3} C^{-1/2} (C^{1/3} n^{-2/3} r)^{3/2} n\right\} \\ &= e^{-4/3 r^{3/2}}, \end{aligned}$$

which agrees the leading order large r asymptotics of the Tracy – Widom GUE distribution [1, Exercise 3.8.3]. Note that the choice of normalizing constant C in this argument is not arbitrary. Taking $C = (1/2)\partial_z^2 g(s, t)|_{z=\zeta}$ is consistent with the normalizing constants needed to see Tracy – Widom GUE limits in, for example, [14, Theorem 1.6] (this is the case $\alpha, \beta \sim \delta_{1/2}$) and in [2, Theorem 1.3]. In the latter case, this was shown to be the constant arising from the KPZ scaling theory in [29]. We also remark that the centering in this argument is likely not correct. As in [12, Theorem 3], we expect that the correct centering should be n times the shape function with α and β given by the empirical distribution of the parameters $\{a_i\}_{i=1}^{\lfloor ns \rfloor}$ and $\{b_j\}_{j=1}^{\lfloor nt \rfloor}$ rather than $ng(s, t)$. This new shape function is not random with respect to $\mathbf{P}_{\mathbf{a}, \mathbf{b}}$ and converges to $g(s, t)$ for almost every realization of the environment. Continuity of the rate function then explains why this difference does not appear at the level of right tail large deviations. The same heuristic suggests that when $E[(a - \underline{\alpha})^{-3}] < \infty$ or $E[(b - \underline{\beta})^{-3}] < \infty$, we should expect KPZ type fluctuations in the critical directions $s/t = c_1$ or $s/t = c_2$, though we do not conjecture the precise limiting distribution in these cases. We also do not address the cases when $E[(a - \underline{\alpha})^{-2}] < \infty$ but $E[(a - \underline{\alpha})^{-3}] = \infty$ or $E[(b - \underline{\beta})^{-2}] < \infty$ but $E[(b - \underline{\beta})^{-3}] = \infty$, though these are interesting questions.

Theorem 2.8. *Suppose that α and β are not both degenerate. For any $s, t > 0$, as $\varepsilon \downarrow 0$,*

$$\mathbb{J}_{s,t}(g(s, t) + \varepsilon) = \begin{cases} \left(-s \mathbb{E}\left[\frac{1}{a - \underline{\alpha}}\right]^2 + t \text{Var}\left[\frac{1}{b + \underline{\alpha}}\right] + t \mathbb{E}\left[\frac{1}{(b + \underline{\alpha})^2}\right] \right)^{-1} \frac{\varepsilon^2}{2} + o(\varepsilon^2) & \text{if } s/t < c_1, \\ \left(s \text{Var}\left[\frac{1}{a + \underline{\zeta}}\right] + t \text{Var}\left[\frac{1}{b - \underline{\zeta}}\right] \right)^{-1} \frac{\varepsilon^2}{2} + o(\varepsilon^2) & \text{if } c_1 \leq s/t \leq c_2, \\ \left(s \text{Var}\left[\frac{1}{a + \underline{\beta}}\right] + s \mathbb{E}\left[\frac{1}{(a + \underline{\beta})^2}\right] - t \mathbb{E}\left[\frac{1}{b - \underline{\beta}}\right]^2 \right)^{-1} \frac{\varepsilon^2}{2} + o(\varepsilon^2) & \text{if } s/t > c_2. \end{cases}$$

We do not have any explicitly computable examples for which the regions $s/t \leq c_1$ and $s/t \geq c_2$ are non-trivial, but we illustrate the results of the last two theorems with a numerical example.

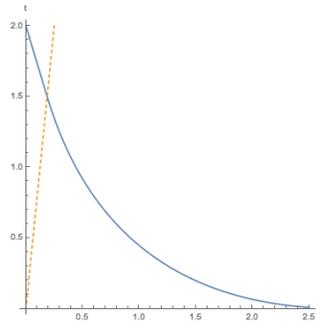


Figure 2: The level set $\{(s, t) : g(s, t) = 1\}$ (solid) and the boundary line $s/t = 1/8$ (dashed).

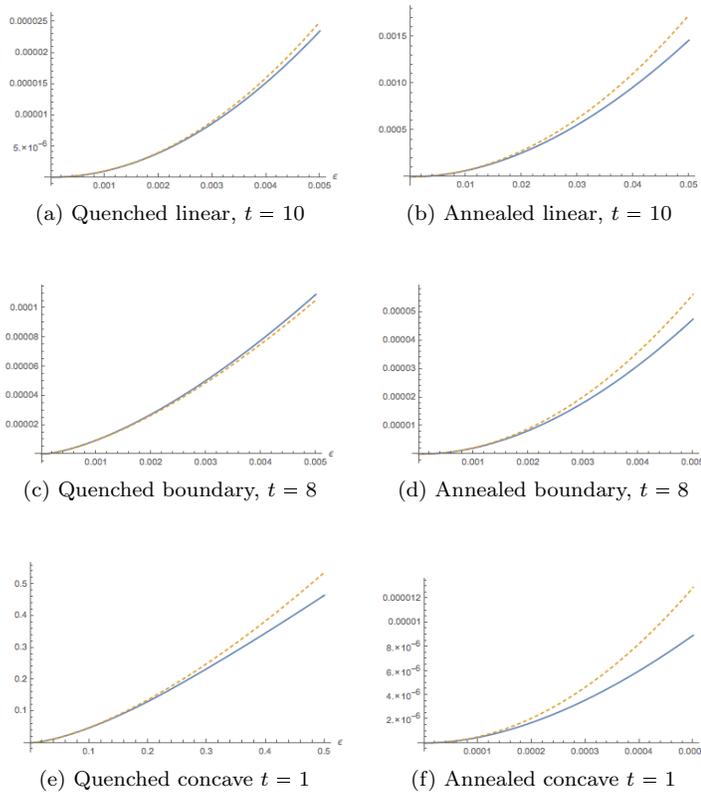


Figure 3: Plot of $\mathbb{J}_{s,t}(g(s, t) + \epsilon)$ and $\mathbf{J}_{s,t}(g(s, t) + \epsilon)$ (solid) and their $\epsilon \downarrow 0$ asymptotics (dashed) with $s = 1$.

Example 2.5. Choose $\alpha = 4(a - 1)^3 1_{[1,2]}(a)da$ and $\beta = \delta_1$. We note that $\alpha = \beta = 1$. Explicit computation shows that $E[(a - 1)^{-2}] = 2$, $E[(b - 1)^{-2}] = \infty$, and $E[(b + 1)^{-2}] = 1/4$. The linear region is then $s/t < 1/8$. This is illustrated in Figure 2.

In Figure 3, we plot numerical approximations of the rate functions against the small ε asymptotics in Theorems 2.7 and 2.8. For example, frame (e) plots $\mathbf{J}_{1,1}(g(1, 1) + \varepsilon)$ against $(4/3)(E(a + \zeta)^{-3} + E(b - \zeta)^{-3})^{-1/2} \varepsilon^{3/2}$, where ζ is the minimizer in (2.3).

3. Variational formulas for the Lyapunov exponents

Note from (2.2) that the probabilities under $\mathbf{P}_{\mathbf{a},\mathbf{b}}^z$ and \mathbb{P}^z of events generated by $\{W(i, 0) : i \in \mathbb{N}\}$ make sense for any $z > -\alpha$. Therefore, we permit ourselves to use notation $\mathbf{P}_{\mathbf{a},\mathbf{b}}^z$ and \mathbb{P}^z (and the corresponding expectations) for $z \geq \underline{\beta}$ and, similarly, for $z \leq -\alpha$ when we work only with $\{W(i, 0) : i \in \mathbb{N}\}$ and $\{W(0, j) : j \in \mathbb{N}\}$, respectively.

Lemma 3.1. *Let $\lambda \in \mathbb{R}$. Suppose that $z > -\alpha$ in (3.1), (3.3), and $z < \underline{\beta}$ in (3.2) and (3.4) below.*

(a) μ -a.s., for any $t > 0$,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left(\lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right) \right] \\ &= \begin{cases} t E \left[\log \left(\frac{a+z}{a+z-\lambda} \right) \right] & \text{if } \lambda \leq \alpha + z, \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{3.1}$$

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left(\lambda \sum_{i=1}^{\lfloor nt \rfloor} W(0, i) \right) \right] \\ &= \begin{cases} t E \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] & \text{if } \lambda \leq \underline{\beta} - z, \\ \infty & \text{otherwise.} \end{cases} \end{aligned} \tag{3.2}$$

(b) For any $t > 0$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}^z \left[\exp \left(\lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right) \right] = \begin{cases} t \log E \left[\frac{a+z}{a+z-\lambda} \right] & \text{if } \lambda \leq \alpha + z, \\ \infty & \text{otherwise.} \end{cases} \tag{3.3}$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}^z \left[\exp \left(\lambda \sum_{i=1}^{\lfloor nt \rfloor} W(0, i) \right) \right] = \begin{cases} t \log E \left[\frac{b-z}{b-z-\lambda} \right] & \text{if } \lambda \leq \underline{\beta} - z, \\ \infty & \text{otherwise.} \end{cases} \tag{3.4}$$

Proof. Using (2.2), we compute

$$\mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] = \begin{cases} \prod_{i=1}^{\lfloor nt \rfloor} \frac{a_i + z}{a_i + z - \lambda} & \text{if } \lambda < \min_{1 \leq i \leq \lfloor nt \rfloor} a_i + z, \\ \infty & \text{otherwise.} \end{cases} \quad (3.5)$$

If $\lambda < \underline{\alpha} + z$ then the first equality in (3.5) holds for all $n \in \mathbb{N}$ μ -a.s and we have

$$\mathbf{E} \left| \log \left(\frac{a + z}{a + z - \lambda} \right) \right| < \infty. \quad (3.6)$$

Hence, by the ergodicity of \mathbf{a} ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} \log \left(\frac{a_i + z}{a_i + z - \lambda} \right) \\ &= t \mathbf{E} \log \left(\frac{a + z}{a + z - \lambda} \right) \quad \mu\text{-a.s.} \end{aligned} \quad (3.7)$$

Moreover, it follows from (3.5) that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] &= \lim_{n \rightarrow \infty} \frac{\lfloor nt \rfloor}{n} \log \mathbf{E} \left[\frac{a + z}{a + z - \lambda} \right] \\ &\rightarrow t \log \mathbf{E} \left[\frac{a + z}{a + z - \lambda} \right]. \end{aligned} \quad (3.8)$$

Next, consider the case $\lambda = \underline{\alpha} + z$. If (3.6) is in force, then both (3.7) and (3.8) still hold. Suppose now that (3.6) fails. By monotonicity,

$$\begin{aligned} \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] &\geq t \mathbf{E} \log \left(\frac{a + z}{a + z - \lambda'} \right) \quad \mu\text{-a.s.}, \\ \liminf_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] &\geq t \log \mathbf{E} \left[\frac{a + z}{a + z - \lambda'} \right] \end{aligned}$$

for any $\lambda' < \lambda$. Letting $\lambda' \uparrow \lambda$ and monotone convergence yield

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] &= \lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor nt \rfloor} W(i, 0) \right\} \right] \\ &= \infty. \end{aligned} \quad (3.9)$$

Finally, consider the case $\lambda > \underline{\alpha} + z$. Then, by the ergodicity of \mathbf{a} , there exists $i \in \mathbb{N}$ such that $\lambda \geq a_i + z$ and the second equality in (3.5) holds for large enough $n \in \mathbb{N}$ μ -a.s. Hence, (3.9).

We have verified (3.1) and (3.3). The proofs of (3.2) and (3.4) are similar. □

Recall the basic properties of the Lyapunov exponents stated in Proposition A.2. For $s, t > 0$ and $\lambda \in \mathbb{R}$, define

$$\mathbf{L}_{s,0}(\lambda) = \lim_{t \downarrow 0} \mathbf{L}_{s,t}(\lambda), \quad \mathbf{L}_{0,t}(\lambda) = \lim_{s \downarrow 0} \mathbf{L}_{s,t}(\lambda),$$

where the limits exist by monotonicity. Define $\mathbb{L}_{s,0}(\lambda)$ and $\mathbb{L}_{0,t}(\lambda)$ similarly. Also, for $k, l \in \mathbb{Z}_+$, let $\theta_{k,l}$ denote the shift given by $\omega(i, j) \mapsto \omega(i + k, j + l)$ for $i, j \in \mathbb{N}$ and $\omega \in \mathbb{R}^{\mathbb{N}^2}$. We next obtain a variational formula involving the Lyapunov exponents.

Lemma 3.2. *Let $z \in (-\underline{\alpha}, \underline{\beta})$ and $\lambda \in (0, \underline{\beta} - z]$. Then*

$$\begin{aligned} & \mathbb{E} \log \left(\frac{a + z + \lambda}{a + z} \right) + \mathbb{E} \log \left(\frac{b - z}{b - z - \lambda} \right) \\ &= \sup_{0 \leq t \leq 1} \left\{ \max \left\{ \mathbf{L}_{t,1}(\lambda) + (1 - t) \mathbb{E} \log \left(\frac{a + z + \lambda}{a + z} \right), \right. \right. \\ & \quad \left. \left. \mathbf{L}_{1,t}(\lambda) + (1 - t) \mathbb{E} \log \left(\frac{b - z}{b - z - \lambda} \right) \right\} \right\}. \end{aligned} \tag{3.10}$$

Also,

$$\begin{aligned} & \log \mathbb{E} \left[\frac{a + z + \lambda}{a + z} \right] + \log \mathbb{E} \left[\frac{b - z}{b - z - \lambda} \right] \\ &= \sup_{0 \leq t \leq 1} \left\{ \max \left\{ \mathbb{L}_{t,1}(\lambda) + (1 - t) \log \mathbb{E} \left[\frac{a + z + \lambda}{a + z} \right], \right. \right. \\ & \quad \left. \left. \mathbb{L}_{1,t}(\lambda) + (1 - t) \log \mathbb{E} \left[\frac{b - z}{b - z - \lambda} \right] \right\} \right\}. \end{aligned} \tag{3.11}$$

Proof of (3.10). We may assume that the left-hand side of (3.10) is finite. (This assumption fails only when $\lambda = \underline{\beta} - z$ and $\mathbb{E} \log(b - \underline{\beta}) = -\infty$ in which case (3.10) clearly holds).

It follows from (1.2) and (2.1) that

$$\begin{aligned} \hat{G}(n, n) &= \max_{1 \leq k \leq n} \left\{ \max \{ G(n - k + 1, n) \circ \theta_{k-1,0} + \hat{G}(k, 0), \right. \\ & \quad \left. G(n, n - k + 1) \circ \theta_{0,k-1} + \hat{G}(0, k) \right\}, \end{aligned}$$

which leads to

$$\begin{aligned} \sum_{1 \leq j \leq n} J(n, j) &= \max_{1 \leq k \leq n} \left\{ \max \left\{ G(n - k + 1, n) \circ \theta_{k-1,0} - \sum_{k < i \leq n} W(i, 0), \right. \right. \\ & \quad \left. \left. G(n, n - k + 1) \circ \theta_{0,k-1} - \sum_{1 \leq i \leq n} W(i, 0) + \sum_{1 \leq j \leq k} W(0, j) \right\} \right\}. \end{aligned} \tag{3.12}$$

Also, note the identity

$$\frac{1}{\mathbf{E}_{\mathbf{a},\mathbf{b}}^z [e^{-\lambda W(i,0)}]} = \frac{a_i + z + \lambda}{a_i + z} = \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} [e^{\lambda W(i,0)}] \text{ for } \lambda > 0 \text{ and } z > -\alpha. \tag{3.13}$$

Using the independence of weights under $\mathbf{P}_{\mathbf{a},\mathbf{b}}^z$, Proposition 2.1, (3.12) and (3.13), we obtain

$$\begin{aligned} & \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i,0) \right\} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0,j) \right\} \right] \\ & \geq \max \left\{ \mathbf{E}_{\tau_{k-1}(\mathbf{a}),\mathbf{b}} [e^{\lambda G(n-k+1,n)}] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq k} W(i,0) \right\} \right], \right. \\ & \quad \left. \mathbf{E}_{\mathbf{a},\tau_{k-1}(\mathbf{b})} [\exp \{ \lambda G(n, n-k+1) \}] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq k} W(0,j) \right\} \right] \right\}. \end{aligned} \tag{3.14}$$

Set $k = \lceil n(1-t) \rceil + 1$ for some $t \in (0, 1)$, apply logarithms to both sides and divide through by n in (3.14). It follows from Proposition A.2 that

$$\begin{aligned} \frac{1}{n} \log \mathbf{E}_{\tau_{k-1}(\mathbf{a}),\mathbf{b}} [e^{\lambda G(n-k+1,n)}] & \rightarrow \mathbf{L}_{t,1}(\lambda), \\ \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\tau_{k-1}(\mathbf{b})} [e^{\lambda G(n,n-k+1)}] & \rightarrow \mathbf{L}_{1,t}(\lambda) \end{aligned}$$

as $n \rightarrow \infty$ along suitable subsequences because (\mathbf{a}, \mathbf{b}) is stationary and \mathbf{L} is deterministic. Hence, also using Lemma 3.1, we obtain

$$\begin{aligned} & \mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right) + \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \\ & \geq \max \left\{ \mathbf{L}_{t,1}(\lambda) + (1-t) \mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right), \mathbf{L}_{1,t}(\lambda) + (1-t) \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \right\}. \end{aligned} \tag{3.15}$$

In particular, \mathbf{L} is finite. By continuity, (3.15) holds with $t = 0$ and $t = 1$ as well.

For the opposite inequality, introduce $L \in \mathbb{N}$ and let $n > L$ such that $\lceil (l+1)n/L \rceil > \lceil ln/L \rceil$ for $0 \leq l < L$. Then, by (3.12) and nonnegativity of the weights,

$$\begin{aligned} & \sum_{1 \leq j \leq n} J(n,j) \\ & \leq \max_{1 \leq l < L} \left\{ \max \left\{ G(\lceil (L-l)n/L \rceil, n) \circ \theta_{\lceil ln/L \rceil, 0} - \sum_{\lceil (l+1)n/L \rceil < i \leq n} W(i,0), \right. \right. \\ & \quad \left. \left. G(n, \lceil (L-l)n/L \rceil) \circ \theta_{0, \lceil ln/L \rceil} - \sum_{1 \leq i \leq n} W(i,0) + \sum_{1 \leq j \leq \lceil (l+1)n/L \rceil} W(0,j) \right\} \right\}, \end{aligned}$$

which implies that

$$\begin{aligned}
 & \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i, 0) \right\} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0, j) \right\} \right] \\
 & \leq \sum_{0 \leq l < L} \mathbf{E}_{\tau_{\lfloor tn/L \rfloor}(\mathbf{a}), \mathbf{b}} \left[e^{\lambda G(\lfloor (L-l)n/L \rfloor, n)} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor (l+1)n/L \rfloor} W(i, 0) \right\} \right] \\
 & \quad + \mathbf{E}_{\mathbf{a}, \tau_{\lfloor tn/L \rfloor}(\mathbf{b})} \left[e^{\lambda G(n, \lfloor (L-l)n/L \rfloor)} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{j=1}^{\lfloor (l+1)n/L \rfloor} W(0, j) \right\} \right].
 \end{aligned} \tag{3.16}$$

Taking logarithms leads to

$$\begin{aligned}
 & \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i, 0) \right\} \right] + \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0, j) \right\} \right] \\
 & \leq \max_{0 \leq l < L} \max \left\{ \log \mathbf{E}_{\tau_{\lfloor tn/L \rfloor}(\mathbf{a}), \mathbf{b}} \left[e^{\lambda G(\lfloor (L-l)n/L \rfloor, n)} \right] \right. \\
 & \quad \left. + \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor (l+1)n/L \rfloor} W(i, 0) \right\} \right], \right. \\
 & \quad \left. \log \mathbf{E}_{\mathbf{a}, \tau_{\lfloor tn/L \rfloor}(\mathbf{b})} \left[e^{\lambda G(n, \lfloor (L-l)n/L \rfloor)} \right] \right. \\
 & \quad \left. + \log \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{j=1}^{\lfloor (l+1)n/L \rfloor} W(0, j) \right\} \right] \right\} + \log(2L).
 \end{aligned}$$

Dividing through by n and letting $n \rightarrow \infty$ along a suitable subsequential limit yield

$$\begin{aligned}
 & \mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right) + \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \\
 & \leq \max_{0 \leq l < L} \max \left\{ \mathbf{L}_{1-l/L, 1}(\lambda) + \frac{l+1}{L} \mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right), \right. \\
 & \quad \left. \mathbf{L}_{1, 1-l/L}(\lambda) + \frac{l+1}{L} \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \right\} \\
 & \leq \sup_{0 \leq t \leq 1} \max \left\{ \mathbf{L}_{t, 1}(\lambda) + (1-t) \mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right), \right. \\
 & \quad \left. \mathbf{L}_{1, t}(\lambda) + (1-t) \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \right\} \\
 & \quad + \frac{1}{L} \left(\mathbf{E} \log \left(\frac{a+z+\lambda}{a+z} \right) + \mathbf{E} \log \left(\frac{b-z}{b-z-\lambda} \right) \right).
 \end{aligned}$$

Letting $L \rightarrow \infty$ completes the proof. □

Proof of (3.11). Some details will be skipped. We may assume that the left-hand side of (3.11) is finite.

Using independence, we can rewrite (3.14) as

$$\begin{aligned}
 & \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{k < i \leq n} W(i, 0) \right\} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0, j) \right\} \right] \\
 & \geq \mathbf{E}_{\tau_{k-1}(\mathbf{a}), \mathbf{b}} \left[e^{\lambda G(n-k+1, n)} \right], \\
 & \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i, 0) \right\} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{k < j \leq n} W(0, j) \right\} \right] \\
 & \geq \mathbf{E}_{\mathbf{a}, \tau_{k-1}(\mathbf{b})} \left[e^{\lambda G(n, n-k+1)} \right].
 \end{aligned} \tag{3.17}$$

The factors on the right-hand side are independent. Applying E yields

$$\begin{aligned}
 & \mathbb{E}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i, 0) \right\} \right] \cdot \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0, j) \right\} \right] \\
 & \geq \max \left\{ \mathbb{E} \left[\exp \{ \lambda G(n-k+1, n) \} \right] \cdot \mathbb{E}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq k} W(i, 0) \right\} \right], \right. \\
 & \quad \left. \mathbb{E} \left[\exp \{ \lambda G(n, n-k+1) \} \right] \cdot \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq k} W(0, j) \right\} \right] \right\},
 \end{aligned} \tag{3.18}$$

where we rearranged terms using that $\{W(i, 0) : i \in \mathbb{N}\}$ and $\{W(0, j) : j \in \mathbb{N}\}$ are both i.i.d. under $\mathbb{P}^{z+\lambda}$ and \mathbb{P}^z . Then, (3.18) leads to \geq half of (3.11) via Proposition A.2 and Lemma 3.1.

For the \leq half of (3.11), suppose that $\lambda < \underline{\beta} - z$ for the moment. Note the inequalities

$$\begin{aligned}
 \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} [e^{\lambda W(i,0)}] &= \frac{a_i + z + \lambda}{a_i + z} \leq \frac{\underline{\alpha} + z + \lambda}{\underline{\alpha} + z}, \\
 \mathbf{E}_{\mathbf{a},\mathbf{b}}^z [e^{\lambda W(0,j)}] &= \frac{b_j - z}{b_j - z - \lambda} \leq \frac{\underline{\beta} - z}{\underline{\beta} - z - \lambda}.
 \end{aligned}$$

It follows from these and (3.16) that

$$\begin{aligned}
 & \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i, 0) \right\} \right] \cdot \mathbf{E}_{\mathbf{a},\mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0, j) \right\} \right] \\
 & \leq \sum_{0 \leq l < L} \left(\frac{\underline{\alpha} + z + \lambda}{\underline{\alpha} + z} \right)^{n/L+1} \mathbf{E}_{\tau_{\lfloor ln/L \rfloor}(\mathbf{a}), \mathbf{b}} \left[e^{\lambda G(\lfloor (L-l)n/L \rfloor, n)} \right] \\
 & \quad \times \mathbf{E}_{\mathbf{a},\mathbf{b}}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor ln/L \rfloor} W(i, 0) \right\} \right] \\
 & \quad + \left(\frac{\underline{\beta} - z}{\underline{\beta} - z - \lambda} \right)^{n/L+1} \mathbf{E}_{\mathbf{a}, \tau_{\lfloor ln/L \rfloor}(\mathbf{b})} \left[e^{\lambda G(n, \lfloor (L-l)n/L \rfloor)} \right]
 \end{aligned}$$

$$\times \mathbf{E}_{\mathbf{a}, \mathbf{b}}^z \left[\exp \left\{ \lambda \sum_{j=1}^{\lfloor ln/L \rfloor} W(0, j) \right\} \right]. \tag{3.19}$$

The point of (3.19) is that the terms on the right-hand side are products of independent factors, which is not the case in (3.16). Applying $\log \mathbb{E}$, we obtain

$$\begin{aligned} & \log \mathbb{E}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{1 \leq i \leq n} W(i, 0) \right\} \right] + \log \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{1 \leq j \leq n} W(0, j) \right\} \right] \\ & \leq \max_{0 \leq l < L} \max \left\{ (n/L + 1) \log \left(\frac{\alpha + z + \lambda}{\alpha + z} \right) + \log \mathbb{E} \left[\exp \{ \lambda G(\lfloor (L-l)n/L \rfloor, n) \} \right] \right. \\ & \quad \left. - \log \mathbb{E}^{z+\lambda} \left[\exp \left\{ \lambda \sum_{i=1}^{\lfloor ln/L \rfloor} W(i, 0) \right\} \right], \right. \\ & \quad (n/L + 1) \log \left(\frac{\beta - z}{\beta - z - \lambda} \right) + \log \mathbb{E} \left[\exp \{ \lambda G(n, \lfloor (L-l)n/L \rfloor) \} \right] \\ & \quad \left. + \log \mathbb{E}^z \left[\exp \left\{ \lambda \sum_{j=1}^{\lfloor ln/L \rfloor} W(0, j) \right\} \right] \right\} \\ & + \log(2L). \end{aligned}$$

Divide through by n and let $n \rightarrow \infty$. If we then send $L \rightarrow \infty$, the result is

$$\begin{aligned} & \log \mathbb{E} \left[\frac{a + z + \lambda}{a + z} \right] + \log \mathbb{E} \left[\frac{b - z}{b - z - \lambda} \right] \\ & \leq \sup_{0 \leq t \leq 1} \left\{ \max \left\{ \mathbb{L}_{t,1}(\lambda) + (1-t) \log \mathbb{E} \left[\frac{a + z + \lambda}{a + z} \right], \right. \right. \\ & \quad \left. \left. \mathbb{L}_{t,1}(\lambda) + (1-t) \log \mathbb{E} \left[\frac{b - z}{b - z - \lambda} \right] \right\} \right\} \end{aligned}$$

for all $\lambda < \beta - z$. The case $\lambda = \beta - z$ also follows because the right-hand side is nondecreasing in λ and the left-hand side, due to monotone convergence, is continuous in λ on $(0, \beta - z]$. \square

Lemma 3.3. For $\lambda > 0$,

$$\begin{aligned} \mathbf{L}_{1,0}(\lambda) &= \mathbb{E} \log \left(\frac{a + \beta}{a + \beta - \lambda} \right), \quad \mathbf{L}_{0,1}(\lambda) = \mathbb{E} \log \left(\frac{b + \alpha}{b + \alpha - \lambda} \right), \quad \text{if } \lambda \leq \alpha + \beta, \\ \mathbf{L}_{1,0}(\lambda) &= \mathbf{L}_{0,1}(\lambda) = \infty \quad \text{otherwise.} \end{aligned} \tag{3.20}$$

$$\begin{aligned} \mathbb{L}_{1,0}(\lambda) &= \log \mathbb{E} \left[\frac{a + \beta}{a + \beta - \lambda} \right], \quad \mathbb{L}_{0,1}(\lambda) = \log \mathbb{E} \left[\frac{b + \alpha}{b + \alpha - \lambda} \right], \quad \text{if } \lambda \leq \alpha + \beta, \\ \mathbb{L}_{1,0}(\lambda) &= \mathbb{L}_{0,1}(\lambda) = \infty \quad \text{otherwise.} \end{aligned} \tag{3.21}$$

Proof. Let $\varepsilon > 0$. On the event $b_1 \leq \underline{\beta} + \varepsilon$, which has positive μ -probability, we have for $n \geq 1/\varepsilon$

$$\begin{aligned} \frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(n, \lfloor n\varepsilon \rfloor)}] &\geq \frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}\left[\exp\left\{\lambda \sum_{1 \leq i \leq n} W(i, 1)\right\}\right] \\ &= \begin{cases} \frac{1}{n} \sum_{i=1}^n \frac{a_i + b_1}{a_i + b_1 - \lambda} & \text{if } \lambda < \min_{1 \leq i \leq n} a_i + b_1, \\ \infty & \text{otherwise} \end{cases} \\ &\geq \begin{cases} \frac{1}{n} \sum_{i=1}^n \frac{a_i + \underline{\beta} + \varepsilon}{a_i + \underline{\beta} + \varepsilon - \lambda} & \text{if } \lambda < \min_{1 \leq i \leq n} a_i + \underline{\beta} + \varepsilon, \\ \infty & \text{otherwise} \end{cases} \\ &= \frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}^{\underline{\beta} + \varepsilon}\left[\exp\left\{\lambda \sum_{1 \leq i \leq n} W(i, 0)\right\}\right]. \end{aligned}$$

Then, by Lemma 3.1,

$$\mathbf{L}_{1, \varepsilon}(\lambda) \geq \begin{cases} \mathbf{E}\left[\log\left(\frac{a + \underline{\beta} + \varepsilon}{a + \underline{\beta} + \varepsilon - \lambda}\right)\right], & \text{if } \lambda \leq \alpha + \underline{\beta} + \varepsilon, \\ \infty & \text{otherwise.} \end{cases}$$

By monotone convergence, letting $\varepsilon \downarrow 0$ yields

$$\mathbf{L}_{1, 0}(\lambda) \geq \begin{cases} \mathbf{E}\left[\log\left(\frac{a + \underline{\beta}}{a + \underline{\beta} - \lambda}\right)\right] & \text{if } \lambda \leq \alpha + \underline{\beta}, \\ \infty & \text{otherwise.} \end{cases}$$

To complete the proof of (3.20), we need

$$\mathbf{L}_{1, 0}(\lambda) \leq \mathbf{E}\left[\log\left(\frac{a + \underline{\beta}}{a + \underline{\beta} - \lambda}\right)\right] \tag{3.22}$$

for $\lambda \in (0, \alpha + \underline{\beta}]$. When $\lambda = \alpha + \underline{\beta}$, we may assume that the right-hand side is finite. Then, $a_i > \alpha$ for $i \in \mathbb{N}$ a.s. and the argument in the paragraph of inequality (3.14) goes through with $z = -\alpha$ as well. Hence,

$$\mathbf{E}\left[\log\left(\frac{a + z + \lambda}{a + z}\right)\right] + \mathbf{E}\left[\log\left(\frac{b - z}{b - z - \lambda}\right)\right] \geq \mathbf{L}_{1, t}(\lambda) + (1 - t) \mathbf{E}\left[\log\left(\frac{b - z}{b - z - \lambda}\right)\right] \tag{3.23}$$

for $t \in [0, 1]$, $z \in [-\alpha, \underline{\beta}]$ and $\lambda \in (0, \underline{\beta} - z]$, which simplifies to

$$\mathbf{E}\left[\log\left(\frac{a + z + \lambda}{a + z}\right)\right] + t \mathbf{E}\left[\log\left(\frac{b - z}{b - z - \lambda}\right)\right] \geq \mathbf{L}_{1, 0}(\lambda). \tag{3.24}$$

Setting $t = 0$ and $z = \underline{\beta} - \lambda$ in (3.24) gives (3.22). The remaining cases are treated similarly. \square

Corollary 3.1. For $s, t > 0$,

$$\begin{aligned} \mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta}) &= s \mathbf{E} \log \left(\frac{a + \underline{\beta}}{a - \underline{\alpha}} \right) + t \mathbf{E} \log \left(\frac{b + \underline{\alpha}}{b - \underline{\beta}} \right). \\ \mathbb{L}_{s,t}(\underline{\alpha} + \underline{\beta}) &= s \log \mathbf{E} \left[\frac{a + \underline{\beta}}{a - \underline{\alpha}} \right] + t \log \mathbf{E} \left[\frac{b + \underline{\alpha}}{b - \underline{\beta}} \right]. \end{aligned}$$

Proof. By concavity and homogeneity,

$$\begin{aligned} \mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta}) &\geq s \mathbf{L}_{1,0}(\underline{\alpha} + \underline{\beta}) + t \mathbf{L}_{0,1}(\underline{\alpha} + \underline{\beta}) \\ &= s \mathbf{E} \left[\log \left(\frac{a + \underline{\beta}}{a - \underline{\alpha}} \right) \right] + t \mathbf{E} \left[\log \left(\frac{b + \underline{\alpha}}{b - \underline{\beta}} \right) \right]. \end{aligned} \tag{3.25}$$

When the right-hand side is finite, the opposite inequality comes from (3.23). $\mathbb{L}_{s,t}(\underline{\alpha} + \underline{\beta})$ is computed similarly. \square

We will use the next lemma to recover the Lyapunov exponents from the variational formulas in Lemma 3.2.

Lemma 3.4. Let $a_0 < b_0$, $A : (a_0, b_0) \rightarrow \mathbb{R}$ be continuous and decreasing, $B : (a_0, b_0) \rightarrow \mathbb{R}$ be continuous and increasing. Let $L : [0, \infty)^2 \rightarrow \mathbb{R}$ be non-decreasing (in each variable), homogeneous, concave and continuous. Assume that

$$A(x) + B(x) = \sup_{0 \leq t \leq 1} \{ \max \{ L(t, 1) + (1 - t)A(x), L(1, t) + (1 - t)B(x) \} \} \tag{3.26}$$

for $a_0 < x < b_0$, $\lim_{x \uparrow b_0} A(x) = \lim_{t \downarrow 0} L(1, t)$ and $\lim_{x \downarrow a_0} B(x) = \lim_{s \downarrow 0} L(s, 1)$. Then

$$L(s, t) = \inf_{a_0 < x < b_0} \{ sA(x) + tB(x) \} \quad \text{for } s, t > 0.$$

Proof. The argument is the same as in [8, Section 5] to prove Theorem 2.1. Assumption (3.26) corresponds to Proposition 4.4 there, and $A(x) = \mathbf{E}[(a + x)^{-1}]$ and $B(x) = \mathbf{E}[(b - x)^{-1}]$ for $x \in (-\underline{\alpha}, \underline{\beta})$. \square

Proof of Theorem 2.1. It follows from Lemma 3.3 that $\mathbf{L}_{s,t}(\lambda) = \infty$ for $\lambda > \underline{\alpha} + \underline{\beta}$. Fix $\lambda \in (0, \underline{\alpha} + \underline{\beta})$ and define

$$\begin{aligned} A(z) &= \mathbf{E} \left[\log \left(\frac{a + z + \lambda}{a + z} \right) \right] \text{ for } z > -\underline{\alpha}, \\ B(z) &= \mathbf{E} \left[\log \left(\frac{b - z}{b - z - \lambda} \right) \right] \text{ for } z < \underline{\beta} - \lambda. \end{aligned}$$

Lemma 3.2 states that

$$A(z) + B(z) = \sup_{0 \leq t \leq 1} \{ \max\{ \mathbf{L}_{t,1}(\lambda) + (1-t)A(z), \mathbf{L}_{1,t}(\lambda) + (1-t)B(z) \} \}$$

for $z \in (-\underline{\alpha}, \underline{\beta} - \lambda)$. Note that A and B are continuous, A is decreasing and B is increasing. Moreover, by Lemma 3.3, $A(\underline{\beta} - \lambda) = \mathbf{L}_{1,0}(\lambda)$ and $B(-\underline{\alpha}) = \mathbf{L}_{0,1}(\lambda)$. Also, $\mathbf{L}_{s,t}(\lambda)$ is finite and, by Proposition A.2, is nondecreasing, homogeneous, concave and continuous. Thus, by Lemma 3.4, $\mathbf{L}_{s,t}(\lambda) = \inf_{-\underline{\alpha} < z < \underline{\beta} - \lambda} \{ sA(z) + tB(z) \}$. The endpoints can be included in the infimum, by monotone convergence. The proof of (2.6) is similar. \square

We close this section with a proof of Theorem 2.2, which is similar to the arguments above.

Proof of Theorem 2.2. We begin with the coupling

$$\begin{aligned} \hat{G}(\lfloor ns \rfloor, \lfloor nt \rfloor) &= \max_{1 \leq k \leq \lfloor ns \rfloor} \left\{ G(\lfloor ns \rfloor - k + 1, \lfloor nt \rfloor) \circ \theta_{k-1,0} + \hat{G}(k, 0) \right\} \\ &\vee \max_{1 \leq k \leq \lfloor nt \rfloor} \left\{ G(\lfloor ns \rfloor, \lfloor nt \rfloor - k + 1) \circ \theta_{0,k-1} + \hat{G}(0, k) \right\}. \end{aligned}$$

Arguing with lim sup and lim inf and coarse graining as above, this leads to the variational problem

$$\begin{aligned} \mathbf{L}_{s,t}^z(\lambda) &= \max_{0 \leq r \leq s} \left\{ \mathbf{L}_{s-r,t}(\lambda) + r \mathbb{E} \left[\log \frac{a+z}{a+z-\lambda} \right] \right\} \\ &\vee \max_{0 \leq u \leq t} \left\{ \mathbf{L}_{s,t-u}(\lambda) + u \mathbb{E} \left[\log \frac{b-z}{b-z-\lambda} \right] \right\}. \end{aligned}$$

Substituting in the variational expression for $\mathbf{L}_{s,t}(\lambda)$, this leads to

$$\begin{aligned} \mathbf{L}_{s,t}^z(\lambda) &= \max_{0 \leq r \leq s} \left\{ \min_{\theta \in [-\underline{\alpha}, \underline{\beta} - \lambda]} \left\{ (s-r) \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right\} \right. \\ &\quad \left. + r \mathbb{E} \left[\log \frac{a+z}{a+z-\lambda} \right] \right\} \\ &\vee \max_{0 \leq u \leq t} \left\{ \min_{\theta \in [-\underline{\alpha}, \underline{\beta} - \lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + (t-u) \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right\} \right. \\ &\quad \left. + u \mathbb{E} \left[\log \frac{b-z}{b-z-\lambda} \right] \right\}. \end{aligned}$$

Applying a minimax theorem (for example [28]), we obtain

$$\mathbf{L}_{s,t}^z(\lambda) = \min_{\theta \in [-\underline{\alpha}, \underline{\beta} - \lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right\}$$

$$\begin{aligned}
 & + \max_{0 \leq r \leq s} r \mathbb{E} \left[\log \frac{(a+z)}{(a+z-\lambda)} \frac{(a+\theta)}{(a+\theta+\lambda)} \right] \Big\} \\
 & \vee \min_{\theta \in [-\alpha, \underline{\beta}-\lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right. \\
 & \left. + \max_{0 \leq u \leq t} u \mathbb{E} \left[\log \frac{(b-z)}{(b-z-\lambda)} \frac{(b-\theta-\lambda)}{(b-\theta)} \right] \right\}.
 \end{aligned}$$

Write $(a+z-\lambda)(a+\theta+\lambda) = (a+z)(a+\theta) + \lambda(z-\theta-\lambda)$ to see that the inner maximum of the first term occurs at $r = s$ if $z-\lambda \leq \theta$ and $r = 0$ if $z-\lambda \geq \theta$. Similarly, $\theta \mapsto (1-\lambda(b-\theta)^{-1})$ is a decreasing function, so the inner maximum of the second term occurs at $u = t$ for $\theta \leq z$ and at $u = 0$ for $\theta \geq z$. Breaking the first minimum over $[-\alpha, \underline{\beta}-\lambda]$ into a minimum over $[-\alpha, z-\lambda]$ and a minimum over $[z-\lambda, \underline{\beta}]$ and the second into a minimum over $[-\alpha, z]$ and a minimum over $[z, \underline{\beta}-\lambda]$, we obtain

$$\begin{aligned}
 & \min_{\theta \in [-\alpha, \underline{\beta}-\lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right. \\
 & \left. + \max_{0 \leq r \leq s} r \mathbb{E} \left[\log \frac{(a+z)}{(a+z-\lambda)} \frac{(a+\theta)}{(a+\theta+\lambda)} \right] \right\} \\
 & = \left\{ s \mathbb{E} \left[\log \frac{a+z}{a+z-\lambda} \right] + t \mathbb{E} \left[\log \frac{b-z+\lambda}{b-z} \right] \right\} \\
 & \quad \wedge \min_{\theta \in [-\alpha, z-\lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right\}
 \end{aligned}$$

and similarly, for the remaining term we have

$$\begin{aligned}
 & \min_{\theta \in [-\alpha, \underline{\beta}-\lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right. \\
 & \left. + \max_{0 \leq u \leq t} u \mathbb{E} \left[\log \frac{(b-z)}{(b-z-\lambda)} \frac{(b-\theta-\lambda)}{(b-\theta)} \right] \right\} \\
 & = \left\{ s \mathbb{E} \left[\log \frac{a+z+\lambda}{a+z} \right] + t \mathbb{E} \left[\log \frac{b-z}{b-z-\lambda} \right] \right\} \\
 & \quad \wedge \min_{\theta \in [z, \underline{\beta}-\lambda]} \left\{ s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right] \right\}.
 \end{aligned}$$

The function

$$\theta \mapsto s \mathbb{E} \left[\log \frac{a+\theta+\lambda}{a+\theta} \right] + t \mathbb{E} \left[\log \frac{b-\theta}{b-\theta-\lambda} \right]$$

is strictly convex with a unique minimizer. Note that the first terms in each of these minima are the values of this function evaluated at $\theta = z-\lambda$ and $\theta = z$. The result follows from strict convexity by considering whether the minimizer lies in $[-\alpha, z]$, $[z, z-\lambda]$, or $[z-\lambda, \underline{\beta}-\lambda]$. \square

4. Extremizers of the variational problems

In this section, we derive some regularity properties of $\mathbf{L}, \mathbb{L}, \mathbf{J}$ and \mathbb{J} by studying the extremizers of their variational representations. The next two lemmas describe the minimizers of (2.5) and (2.6). See Figure 4 for an illustration.

Lemma 4.1. Fix $s, t > 0$ and define $F = F(z, \lambda)$ for $0 < \lambda < \underline{\alpha} + \underline{\beta}$ and $-\underline{\alpha} \leq z \leq \underline{\beta} - \lambda$ by

$$F(z, \lambda) = s \mathbb{E} \log \left(\frac{a + z + \lambda}{a + z} \right) + t \mathbb{E} \log \left(\frac{b - z}{b - z - \lambda} \right). \tag{4.1}$$

For each $\lambda \in (0, \underline{\alpha} + \underline{\beta})$, there exists a unique $z_\star = z_\star(\lambda) \in [-\underline{\alpha}, \underline{\beta} - \lambda]$ such that $\mathbf{L}_{s,t}(\lambda) = F(z_\star, \lambda)$. We have $z_\star = -\underline{\alpha}$ if and only if

$$-s \mathbb{E} \left[\frac{1}{(a - \underline{\alpha} + \lambda)(a - \underline{\alpha})} \right] + t \mathbb{E} \left[\frac{1}{(b + \underline{\alpha} - \lambda)(b + \underline{\alpha})} \right] \geq 0, \tag{4.2}$$

and $z_\star = \underline{\beta} - \lambda$ if and only if

$$-s \mathbb{E} \left[\frac{1}{(a + \underline{\beta})(a + \underline{\beta} - \lambda)} \right] + t \mathbb{E} \left[\frac{1}{(b - \underline{\beta})(b - \underline{\beta} + \lambda)} \right] \leq 0. \tag{4.3}$$

Define $\lambda_1 = \inf\{\lambda \in (0, \underline{\alpha} + \underline{\beta}) : (4.2) \text{ holds.}\} \wedge (\underline{\alpha} + \underline{\beta})$ and $\lambda_2 = \inf\{\lambda \in (0, \underline{\alpha} + \underline{\beta}) : (4.3) \text{ holds.}\} \wedge (\underline{\alpha} + \underline{\beta})$. Then $z_\star = -\underline{\alpha}$ if and only if $\lambda \geq \lambda_1$, and $z_\star = \underline{\beta} - \lambda$ if and only if $\lambda \geq \lambda_2$. For $0 < \lambda < \lambda_0 = \lambda_1 \wedge \lambda_2$, we have $\partial_z F(z_\star, \lambda) = 0$. Moreover, z_\star is continuous on $(0, \underline{\alpha} + \underline{\beta})$ and continuously differentiable on $(0, \underline{\alpha} + \underline{\beta}) \setminus \{\lambda_0\}$. We have $-1 < z_\star' < 0$ for $0 < \lambda < \lambda_0$, $\lim_{\lambda \downarrow 0} z_\star = \zeta(s, t)$ and $\lim_{\lambda \uparrow \underline{\alpha} + \underline{\beta}} z_\star = -\underline{\alpha}$.

Lemma 4.2. Lemma 4.1 holds verbatim if $\mathbf{L}_{s,t}$, (4.1), (4.2) and (4.3) are replaced with $\mathbb{L}_{s,t}$,

$$F(z, \lambda) = s \log \mathbb{E} \left[\frac{a + z + \lambda}{a + z} \right] + t \log \mathbb{E} \left[\frac{b - z}{b - z - \lambda} \right], \tag{4.4}$$

$$-s \frac{\mathbb{E} \left[\frac{1}{(a - \underline{\alpha})^2} \right]}{\mathbb{E} \left[\frac{a - \underline{\alpha} + \lambda}{a - \underline{\alpha}} \right]} + t \frac{\mathbb{E} \left[\frac{1}{(b + \underline{\alpha} - \lambda)^2} \right]}{\mathbb{E} \left[\frac{b + \underline{\alpha}}{b + \underline{\alpha} - \lambda} \right]} \geq 0, \tag{4.5}$$

$$-s \frac{\mathbb{E} \left[\frac{1}{(a + \underline{\beta} - \lambda)^2} \right]}{\mathbb{E} \left[\frac{a + \underline{\beta}}{a + \underline{\beta} - \lambda} \right]} + t \frac{\mathbb{E} \left[\frac{1}{(b - \underline{\beta})^2} \right]}{\mathbb{E} \left[\frac{b - \underline{\beta} + \lambda}{b - \underline{\beta}} \right]} \leq 0, \tag{4.6}$$

respectively. Here, the left-hand sides of (4.5) and (4.6) are interpreted as $-\infty$ and ∞ when $\mathbb{E}[(a - \underline{\alpha})^{-1}] = \infty$ and $\mathbb{E}[(b - \underline{\beta})^{-1}] = \infty$, respectively.

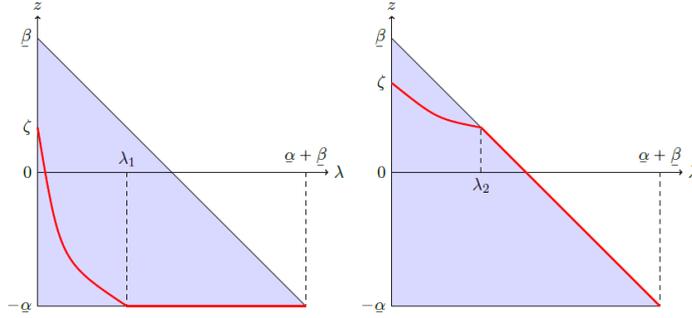


Figure 4: Sketches of the graph of the minimizers in (2.5) and (2.6) assuming (4.2) and (4.5), respectively (left) and assuming (4.3) and (4.6), respectively (right).

Proof of Lemma 4.1. Since $\partial_z^2 F > 0$, the existence and the uniqueness of z_\star follows. Also, $z_\star = -\underline{\alpha}$ if and only if $\partial_z F(-\underline{\alpha}, \lambda) \geq 0$, which is (4.2). We note that $\partial_z F(-\underline{\alpha}, \lambda) = -\infty$ if $E[(a - \underline{\alpha})^{-1}] = \infty$ and, otherwise, $\lambda^{-1} \partial_z F(-\underline{\alpha}, \lambda)$ is a continuous, increasing function of $\lambda \in (0, \underline{\alpha} + \underline{\beta})$. Therefore, $z_\star = -\underline{\alpha}$ if and only if $\lambda \geq \lambda_1$. We similarly observe (4.3) and the equivalence of $z_\star = \underline{\beta} - \lambda$ and $\lambda \geq \lambda_2$. (Because $\partial_z F$ is increasing in z , we cannot have λ_1 and λ_2 both less than $\underline{\alpha} + \underline{\beta}$).

When $\lambda < \lambda_0$, the minimizer is the unique $z_\star \in (-\underline{\alpha}, \underline{\beta} - \lambda)$ satisfying

$$\partial_z F(z_\star, \lambda) = 0. \tag{4.7}$$

By the implicit function theorem, z_\star is continuously differentiable for $0 < \lambda < \lambda_0$ with derivative

$$z_\star'(\lambda) = -\frac{\partial_\lambda \partial_z F(z_\star, \lambda)}{\partial_z^2 F(z_\star, \lambda)}. \tag{4.8}$$

Observing that

$$\begin{aligned} \partial_\lambda \partial_z F(z_\star, \lambda) &> -s E\left[\frac{1}{(a + z_\star)(a + z_\star + \lambda)}\right] + t E\left[\frac{1}{(b - z_\star)(b - z_\star - \lambda)}\right] \\ &= \lambda^{-1} \partial_z F(z_\star, \lambda) = 0, \\ \partial_z^2 F(z_\star, \lambda) - \partial_\lambda \partial_z F(z_\star, \lambda) &= s E\left[\frac{1}{(a + z_\star)^2}\right] - t E\left[\frac{1}{(b - z_\star)^2}\right] \\ &> s E\left[\frac{1}{(a + z_\star)(a + z_\star + \lambda)}\right] - t E\left[\frac{1}{(b - z_\star)(b - z_\star - \lambda)}\right] = 0, \end{aligned}$$

we conclude that $-1 < z_\star'(\lambda) < 0$. In particular, z_\star is monotone and has limits as $\lambda \downarrow 0$ and $\lambda \uparrow \lambda_0$. We also have continuous differentiability of z_\star for $\lambda > \lambda_0$. Now, supposing $\lambda_0 \in (0, \underline{\alpha} + \underline{\beta})$, we show that z_\star is continuous at λ_0 . Letting $\lambda \uparrow \lambda_0$ in (4.7), we obtain

$$\partial_z F(\lim_{\lambda \uparrow \lambda_0} z_\star(\lambda), \lambda_0) = 0. \tag{4.9}$$

Since the minimizer occurs at the boundary when $\lambda = \lambda_0$, we deduce from (4.9) that $\lim_{\lambda \uparrow \lambda_1} z_\star(\lambda) = -\underline{\alpha}$ and $\lim_{\lambda \uparrow \lambda_2} z_\star(\lambda) = \underline{\beta} - \lambda_2$ when $\lambda_0 = \lambda_1$ and $\lambda_0 = \lambda_2$, respectively.

Since $z_\star(\lambda) \in [-\underline{\alpha}, \underline{\beta} - \lambda]$, we have $\lim_{\lambda \uparrow \underline{\alpha} + \underline{\beta}} z_\star(\lambda) = -\underline{\alpha}$. Set $z_\star(0) = \lim_{\lambda \downarrow 0} z_\star(\lambda)$. To calculate this limit, we consider several cases. If $\lambda_0 > 0$ then we can let $\lambda \downarrow 0$ in (4.7) and obtain

$$0 = \partial_z F(z_\star(0), 0) = -s \mathbb{E}\left[\frac{1}{(a + z_\star(0))^2}\right] + t \mathbb{E}\left[\frac{1}{(b - z_\star(0))^2}\right] = \partial_z g_{z_\star(0)}(s, t),$$

which implies $z_\star(0) = \zeta$. If $\lambda_1 = 0$ then $\partial_z F(-\underline{\alpha}, 0) = \partial_z g_{-\underline{\alpha}}(s, t) \geq 0$ and if $\lambda_2 = 0$ then $\partial_z F(\underline{\beta}, 0) = \partial_z g_{\underline{\beta}}(s, t) \leq 0$. Hence, we get $\zeta = -\underline{\alpha} = z_\star(0)$ and $\zeta = \underline{\beta} = z_\star(0)$, respectively. \square

We omit the proof of Lemma 4.2 which is similar to that of Lemma 4.1.

Lemma 4.3. *For each $s, t > 0$, $\mathbf{L}_{s,t}$ is continuously differentiable on $[0, \underline{\alpha} + \underline{\beta}]$ and $\mathbf{L}'_{s,t}(0) = g(s, t)$. Furthermore, $\mathbf{L}'_{s,t}$ is continuously differentiable on $(0, \underline{\alpha} + \underline{\beta}) \setminus \{\lambda_0\}$ and $\mathbf{L}''_{s,t} > 0$. The same statements also hold for $\mathbb{L}_{s,t}$.*

Proof. Let us write L for $\mathbf{L}_{s,t}$ and $F = F(z, \lambda)$ be given by (4.4). Using Lemma 4.1, we compute

$$L'(\lambda) = \partial_z F(z_\star, \lambda) z_\star'(\lambda) + \partial_\lambda F(z_\star, \lambda) = s \mathbb{E}\left[\frac{1}{a + z_\star + \lambda}\right] + t \mathbb{E}\left[\frac{1}{b - z_\star - \lambda}\right] \tag{4.10}$$

for $0 < \lambda < \lambda_0$. Differentiating again, we obtain

$$\begin{aligned} L''(\lambda) &= \partial_z \partial_\lambda F(z_\star, \lambda) z_\star'(\lambda) + \partial_\lambda^2 F(z_\star, \lambda) \\ &= \frac{\partial_z^2 F(z_\star, \lambda) \partial_\lambda^2 F(z_\star, \lambda) - \partial_z \partial_\lambda F(z_\star, \lambda)^2}{\partial_z^2 F(z_\star, \lambda)} > 0, \end{aligned}$$

where the inequality comes from $\partial_z^2 F(z_\star, \lambda) > \partial_\lambda \partial_z F(z_\star, \lambda)$ and $\partial_\lambda^2 F = \partial_\lambda \partial_z F$. For $\lambda > \lambda_1$,

$$L'(\lambda) = s \mathbb{E}\left[\frac{1}{a - \underline{\alpha} + \lambda}\right] + t \mathbb{E}\left[\frac{1}{b + \underline{\alpha} - \lambda}\right], \tag{4.11}$$

$$L''(\lambda) = -s \mathbb{E}\left[\frac{1}{(a - \underline{\alpha} + \lambda)^2}\right] + t \mathbb{E}\left[\frac{1}{(b + \underline{\alpha} - \lambda)^2}\right] > \partial_z F(-\underline{\alpha}, \lambda) > 0. \quad (4.12)$$

Also, for $\lambda > \lambda_2$,

$$L'(\lambda) = s \mathbb{E}\left[\frac{1}{a + \underline{\beta} - \lambda}\right] + t \mathbb{E}\left[\frac{1}{b - \underline{\beta} + \lambda}\right], \quad (4.13)$$

$$L''(\lambda) = s \mathbb{E}\left[\frac{1}{(a + \underline{\beta} - \lambda)^2}\right] - t \mathbb{E}\left[\frac{1}{(b - \underline{\beta} + \lambda)^2}\right] > -\partial_z F(\underline{\beta} - \lambda, \lambda) > 0. \quad (4.14)$$

We have verified that L is continuously differentiable on $(0, \underline{\alpha} + \underline{\beta}) \setminus \{\lambda_0\}$ and L' is increasing.

We next note that L is also continuously differentiable at λ_0 when $\lambda_0 \in (0, \underline{\alpha} + \underline{\beta})$, for which it suffices to check that the left and right limits of L' at λ_0 match. First, we consider the case $\lambda_1 \in (0, \underline{\alpha} + \underline{\beta})$. Then, as $\lambda \uparrow \lambda_1$, (4.10) tends to $s \mathbb{E}[(a - \underline{\alpha} + \lambda_1)^{-1}] + t \mathbb{E}[(b - \underline{\alpha} - \lambda_1)^{-1}]$, which equals the $\lambda \downarrow \lambda_1$ limit of (4.11). Now, suppose that $\lambda_2 \in (0, \underline{\alpha} + \underline{\beta})$. Then, as $\lambda \uparrow \lambda_2$, (4.10) tends to $s \mathbb{E}[(a + \underline{\beta})^{-1}] + t \mathbb{E}[(b - \underline{\beta})^{-1}]$, which is the same as

$$\begin{aligned} & s \mathbb{E}\left[\frac{1}{a + \underline{\beta} - \lambda_2}\right] + t \mathbb{E}\left[\frac{1}{b - \underline{\beta} + \lambda_2}\right] + \partial_z F(\underline{\beta} - \lambda_2, \lambda_2) \\ &= s \mathbb{E}\left[\frac{1}{a + \underline{\beta} - \lambda_2}\right] + t \mathbb{E}\left[\frac{1}{b - \underline{\beta} + \lambda_2}\right], \end{aligned}$$

the $\lambda \downarrow \lambda_2$ limit of (4.13).

We next calculate $L'(0) = \lim_{\lambda \downarrow 0} L'(\lambda)$. If $\lambda_0 > 0$ then $\lambda \downarrow 0$ limit of (4.10) gives

$$L'(0) = s \mathbb{E}\left[\frac{1}{a + \zeta}\right] + t \mathbb{E}\left[\frac{1}{b - \zeta}\right] = g(s, t).$$

In the cases $\lambda_1 = 0$ and λ_2 then $\zeta = -\underline{\alpha}$ and $\zeta = \underline{\beta}$, respectively. Hence, letting $\lambda \downarrow 0$ in (4.11) and (4.13), respectively, we still obtain $L'(0) = g(s, t)$.

The asserted properties of \mathbb{L} are proved similarly. □

Since $\mathbf{L}'_{s,t}$ increasing, $\mathbf{L}'_{s,t}(\lambda)$ has a limit (possibly ∞) as $\lambda \uparrow \underline{\alpha} + \underline{\beta}$, which we denote by $\mathbf{L}'_{s,t}(\underline{\alpha} + \underline{\beta})$. Similarly, let us write $\mathbb{L}'_{s,t}(\underline{\alpha} + \underline{\beta})$ for $\lim_{\lambda \uparrow \underline{\alpha} + \underline{\beta}} \mathbb{L}'_{s,t}(\lambda)$. The precise values of these limits will be needed in the next section.

Corollary 4.1. *Fix $s, t > 0$.*

$$\mathbf{L}'_{s,t}(\underline{\alpha} + \underline{\beta}) = \begin{cases} s \mathbb{E}\left[\frac{1}{a - \underline{\alpha}}\right] + t \mathbb{E}\left[\frac{1}{b + \underline{\alpha}}\right] \\ \text{if } -s \mathbb{E}\left[\frac{1}{(a - \underline{\alpha})(a + \underline{\beta})}\right] + t \mathbb{E}\left[\frac{1}{(b + \underline{\alpha})(b - \underline{\beta})}\right] \leq 0, \\ s \mathbb{E}\left[\frac{1}{a + \underline{\beta}}\right] + t \mathbb{E}\left[\frac{1}{b - \underline{\beta}}\right] \quad \text{otherwise.} \end{cases}$$

$$\mathbb{L}'_{s,t}(\underline{\alpha} + \underline{\beta}) = \begin{cases} s \frac{\mathbb{E}\left[\frac{a + \underline{\beta}}{(a - \underline{\alpha})^2}\right]}{\mathbb{E}\left[\frac{a + \underline{\beta}}{a - \underline{\alpha}}\right]} + t \frac{\mathbb{E}\left[\frac{1}{b - \underline{\beta}}\right]}{\mathbb{E}\left[\frac{b + \underline{\alpha}}{b - \underline{\beta}}\right]} \\ \text{if } -s \frac{\mathbb{E}\left[\frac{1}{(a - \underline{\alpha})^2}\right]}{\mathbb{E}\left[\frac{a + \underline{\beta}}{a - \underline{\alpha}}\right]} + t \frac{\mathbb{E}\left[\frac{1}{(b - \underline{\beta})^2}\right]}{\mathbb{E}\left[\frac{b + \underline{\alpha}}{b - \underline{\beta}}\right]} \leq 0, \\ s \frac{\mathbb{E}\left[\frac{1}{a - \underline{\alpha}}\right]}{\mathbb{E}\left[\frac{a + \underline{\beta}}{a - \underline{\alpha}}\right]} + t \frac{\mathbb{E}\left[\frac{b + \underline{\alpha}}{(b - \underline{\beta})^2}\right]}{\mathbb{E}\left[\frac{b + \underline{\alpha}}{b - \underline{\beta}}\right]} \quad \text{otherwise.} \end{cases}$$

The next lemma establishes continuous differentiability of $\mathbf{J}_{s,t}(r)$ and $\mathbb{J}_{s,t}(r)$ and shows that these functions are linear in r for $r > \mathbf{L}'_{s,t}(\underline{\alpha} + \underline{\beta})$ and $r > \mathbb{L}'_{s,t}(\underline{\alpha} + \underline{\beta})$, respectively.

Lemma 4.4. Fix $s, t > 0$. For each $r \geq g(s, t)$, there exists a unique $\lambda_*(r) \in [0, \underline{\alpha} + \underline{\beta}]$ such that $\mathbf{J}_{s,t}(r) = \lambda_* r - \mathbf{L}_{s,t}(\lambda_*)$. Moreover, $\mathbf{J}_{s,t}$ is continuously differentiable and $\mathbf{J}'_{s,t}(r) = \lambda_*(r)$ for $r \geq g(s, t)$. If $r > g(s, t)$, then $\lambda_* > 0$. If $r \geq \mathbf{L}'_{s,t}(\underline{\alpha} + \underline{\beta})$ then $\lambda_* = \underline{\alpha} + \underline{\beta}$, while if $r \in [g(s, t), \mathbf{L}'_{s,t}(\underline{\alpha} + \underline{\beta})]$ then $\mathbf{L}'_{s,t}(\lambda_*) = r$. The same statements hold if we replace $\mathbf{J}_{s,t}$ and $\mathbf{L}_{s,t}$ with $\mathbb{J}_{s,t}$ and $\mathbb{L}_{s,t}$, respectively.

Proof. We have $J(r) = \sup_{0 < \lambda < \underline{\alpha} + \underline{\beta}} \{\lambda r - L(\lambda)\}$, where (L, J) pair refers to either $(\mathbf{L}_{s,t}, \mathbf{J}_{s,t})$ or $(\mathbb{L}_{s,t}, \mathbb{J}_{s,t})$. The λ -derivative of the function inside the supremum is $r - L'(\lambda)$. By Lemma 4.3, L' is continuous and increasing from $g(s, t)$ to the limit $L'(\underline{\alpha} + \underline{\beta})$ on $(0, \underline{\alpha} + \underline{\beta})$. It follows that the unique maximizer λ_* is at $\underline{\alpha} + \underline{\beta}$ if $r \geq L'(\underline{\alpha} + \underline{\beta})$ and at $(L')^{-1}(r)$, otherwise. In addition, λ_* is increasing and continuous on $[g(s, t), +\infty)$. Since L' is differentiable and has nonzero derivative for $\lambda \in (0, \underline{\alpha} + \underline{\beta}) \setminus \lambda_0$, whenever $r \neq L'(\lambda_0)$, we have $J'(r) = \lambda_*(r) + \lambda_*'(r)r - L'(\lambda_*)\lambda_*'(r) = \lambda_*(r)$. Then continuity of λ_* implies that J is continuously differentiable for all $r \geq g(s, t)$ including $L'(\lambda_0)$ when $\lambda_0 \in (0, \underline{\alpha} + \underline{\beta})$. □

Proof of Theorem 2.6. This theorem is included in the preceding lemma. □

5. Left tail estimates

We now estimate the left tail in both the quenched and annealed settings. The first result shows that in the quenched case, the rate n large deviation rate function will be trivial for deviations to the left of the shape function $g(s, t)$.

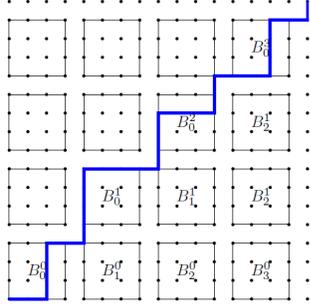


Figure 5: A path passing through the bottom-left and top-right vertices of B_0^j for each j .

This proof is based on the proof of [22, Theorem 4.1], which was adapted from an argument in [16].

Proof of Lemma 2.1. First, fix $s, t, \varepsilon > 0$ and rational. Take $m \in \mathbb{N}$ large enough that $m^{-1}\mathbb{E} G(\lfloor ms \rfloor, \lfloor mt \rfloor) \geq g(s, t) - \varepsilon/2$. We coarse grain the lattice into pairwise disjoint translates of the set $\{1, \dots, \lfloor ms \rfloor\} \times \{1, \dots, \lfloor mt \rfloor\}$. Toward this end, define

$$A_{a,b}^{k,\ell} = \{1 + a, \dots, a + k\} \times \{1 + b, \dots, \ell + b\}, \quad B_i^j = A_{(j+i)\lfloor ms \rfloor, j\lfloor mt \rfloor}^{\lfloor mx \rfloor, \lfloor my \rfloor}.$$

Take n large and let $L = \lfloor n/m - \lfloor \sqrt{n} \rfloor - 2 \rfloor$. For each such $k \leq \lfloor \sqrt{n} \rfloor$, define a diagonal by $D_k = \bigcup_{j=0}^L B_k^j$. We observe that the passage time from the bottom left corner of B_i^j to the top-right corner of B_i^j , $G_{i,j} \equiv G(\lfloor ms \rfloor, \lfloor mt \rfloor) \circ \tau_{(i+j)\lfloor ms \rfloor, j\lfloor mt \rfloor}$, has the same distribution as $G_{0,0}$ under \mathbb{P} . Moreover, if $(i_1, j_1) \neq (i_2, j_2)$, then $B_{i_1}^{j_1} \cap B_{i_2}^{j_2} = \emptyset$ and consequently $\{G_{i,j}\}_{i,j \geq 0}$ forms an independent family under $\mathbf{P}_{\mathbf{a},\mathbf{b}}$.

Denote by Π_k the collection of paths from $(1,1)$ to $(\lfloor ns \rfloor, \lfloor nt \rfloor)$ passing through the bottom-left and top-right vertices of B_k^j for each j . See Figure 5. We have

$$G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq \max_{k \leq \lfloor \sqrt{n} \rfloor} \max_{\pi \in \Pi_k} \sum_{(i,j) \in \pi} W(i,j) \geq \max_{k \leq \lfloor \sqrt{n} \rfloor} \sum_{j \leq L} G_{k,j}.$$

It follows that

$$\begin{aligned} \mathbf{P}_{\mathbf{a},\mathbf{b}}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \leq (g(s,t) - \varepsilon)) &\leq \mathbf{P}_{\mathbf{a},\mathbf{b}}\left(\max_{k \leq \lfloor \sqrt{n} \rfloor} n^{-1} \sum_{j=0}^L G_{k,j} \leq g(s,t) - \varepsilon\right) \\ &= \prod_{k=0}^{\lfloor \sqrt{n} \rfloor} \mathbf{P}_{\mathbf{a},\mathbf{b}}\left(n^{-1} \sum_{j=0}^L G_{k,j} \leq g(s,t) - \varepsilon\right). \end{aligned}$$

Now, fix $\lambda > 0$ sufficiently small that $C \equiv \lambda m \varepsilon / 2 - (\lambda^2 / 2) \mathbb{E} G_{0,0}^2 > 0$ and $\lambda \mathbb{E} G_{0,0} - (\lambda^2 / 2) \mathbb{E} G_{0,0}^2 < 1$ and notice that

$$\mathbf{E}_{\mathbf{a},\mathbf{b}} [\exp\{-\lambda G_{j,k}\}] = \mathbf{E}_{\mathbf{a},\mathbf{b}} [\exp\{-\lambda G_{0,0}\}] \circ \tau_{(j+k)\lfloor ms\rfloor, k\lfloor mt\rfloor}.$$

The ergodic theorem then implies that the following limit holds μ almost surely:

$$\lim_{L \rightarrow \infty} \frac{1}{L} \sum_{j=0}^L \log \mathbf{E}_{\mathbf{a},\mathbf{b}} [\exp\{-\lambda G_{k,j}\}] = \mathbb{E} [\log \mathbf{E}_{\mathbf{a},\mathbf{b}} [\exp\{-\lambda G_{0,0}\}]].$$

Jensen’s inequality gives $\mathbb{E} [\log \mathbf{E}_{\mathbf{a},\mathbf{b}} [\exp\{-\lambda G_{0,0}\}]] \leq \log \mathbb{E} [\exp\{-\lambda G_{0,0}\}] < -\lambda \mathbb{E} G_{0,0} + (\lambda^2 / 2) \mathbb{E} G_{0,0}^2$. By the exponential Markov inequality and independence under $\mathbf{P}_{\mathbf{a},\mathbf{b}}$, we have

$$\begin{aligned} & \frac{1}{L} \log \mathbf{P}_{\mathbf{a},\mathbf{b}} \left(\sum_{j=0}^L G_{k,j} < n(g(s, t) - \varepsilon) \right) \\ & \leq \frac{1}{L} \left(\sum_{j=0}^L \log \mathbf{E}_{\mathbf{a},\mathbf{b}} [\exp\{-\lambda G_{j,k}\}] + \lambda n(g(s, t) - \varepsilon) \right). \end{aligned}$$

Recalling that $L^{-1}n \rightarrow m$ as $n \rightarrow \infty$, and our assumption that $\mathbb{E} G_{0,0} > m(g(s, t) - \varepsilon/2)$, it follows that

$$\limsup_{L \rightarrow \infty} L^{-1} \log \mathbf{P}_{\mathbf{a},\mathbf{b}} \left(\sum_{j=0}^L G_{k,j} < n(g(s, t) - \varepsilon) \right) \leq -\lambda m \frac{\varepsilon}{2} + \frac{\lambda^2}{2} \mathbb{E} G_{0,0}^2 = -C$$

almost surely. Therefore, for each k there exists a random N_k so that for $n \geq N_k$

$$\mathbf{P}_{\mathbf{a},\mathbf{b}} \left(\sum_{j=0}^L G_{k,j} < n(g(s, t) - \varepsilon) \right) \leq \exp\left\{-n \frac{C}{2m}\right\}.$$

For any fixed K and $n \geq \max_{k \leq K} N_k$, we see that \mathbb{P} almost surely we have

$$\begin{aligned} & -\frac{1}{n} \log \mathbf{P}_{\mathbf{a},\mathbf{b}} (n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \leq (g(s, t) - \varepsilon)) \\ & \geq \sum_{k=0}^K -\frac{1}{n} \log \mathbf{P}_{\mathbf{a},\mathbf{b}} \left(n^{-1} \sum_{j=0}^L G_{k,j} \leq g(s, t) - \varepsilon \right) \\ & \geq K \frac{C}{2m}. \end{aligned}$$

Sending $n \rightarrow \infty$ and then $K \rightarrow \infty$ gives the result for fixed $s, t, \varepsilon > 0$. For the general result, we work on the μ almost sure set where the result holds simultaneously for all rational $s, t, \varepsilon > 0$. Take $s, t, \varepsilon > 0$ and $s_1 < s$ and $t_1 < t$

rational with the property that $\varepsilon - g(s, t) + g(s_1, t_1) > \varepsilon_1 > 0$ for rational ε_1 . This is possible by continuity of g . The result follows from observing that

$$\mathbf{P}_{\mathbf{a}, \mathbf{b}}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \leq g(s, t) - \varepsilon) \leq \mathbf{P}_{\mathbf{a}, \mathbf{b}}(n^{-1}G(\lfloor ns_1 \rfloor, \lfloor nt_1 \rfloor) \leq g(s_1, t_1) - \varepsilon_1).$$

□

Corollary 5.1. $\lim_{n \rightarrow \infty} n^{-1} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}[\exp\{-\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)\}] = -\lambda g(s, t)$ μ -a.s. for $s, t, \lambda > 0$.

Essentially the same argument as in Lemma 2.1 restricted to a single diagonal D_0 (so that the last passage times on B_0^j are i.i.d. under \mathbb{P}) shows that for $r \in (0, g(s, t))$, we have

$$\liminf_{n \rightarrow \infty} -n^{-1} \log \mathbb{P}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \leq r) > 0.$$

To show that n is the correct rate for certain left tail large deviations, we need to show that the corresponding limsup is finite for some $r \in (0, g(s, t))$. We begin by considering the natural mechanism for these deviations, which we stated previously in Section 2 as Lemma 2.2.

Proof of Lemma 2.2. We may assume without loss of generality that $\{\nu_1 \in \mathcal{M}^\alpha, \nu_2 \in \mathcal{M}^\beta : g_{\nu_1, \nu_2}(s, t) \in (x, y)\} \neq \emptyset$ since the right hand side is infinite otherwise. Fix a pair ν_1, ν_2 from this set and introduce the notation

$$A_n = \{n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \in (x, y)\}, \quad \frac{d\nu_1}{d\alpha}(a) = \varphi(a), \quad \frac{d\nu_2}{d\beta}(b) = \psi(b).$$

Since A_n is measurable with respect to $\sigma(W(i, j) : 1 \leq i \leq \lfloor ns \rfloor, 1 \leq j \leq \lfloor nt \rfloor)$, we see that

$$\begin{aligned} \mathbb{P}_{\alpha, \beta}(A_n) &= \mathbf{E}_{\alpha, \beta}[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n)] \geq \mathbf{E}_{\alpha, \beta} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n) \prod_{i=1}^{\lfloor ns \rfloor} 1_{\{\varphi(a_i) > 0\}} \prod_{j=1}^{\lfloor nt \rfloor} 1_{\{\psi(b_j) > 0\}} \right] \\ &= \mathbf{E}_{\nu_1, \nu_2} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n) \prod_{i=1}^{\lfloor ns \rfloor} \varphi(a_i)^{-1} \prod_{j=1}^{\lfloor nt \rfloor} \psi(b_j)^{-1} \right]. \end{aligned}$$

Taking logs and applying Jensen’s inequality shows that

$$\begin{aligned} &-\frac{1}{n} \log \mathbb{P}_{\alpha, \beta}(A_n) \\ &\leq -\frac{1}{n} \log \mathbf{E}_{\nu_1, \nu_2} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n) \prod_{i=1}^{\lfloor ns \rfloor} \varphi(a_i)^{-1} \prod_{j=1}^{\lfloor nt \rfloor} \psi(b_j)^{-1} \right] \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{n \mathbb{P}_{\nu_1, \nu_2}(A_n)} \mathbb{E}_{\nu_1, \nu_2} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n) \left(\sum_{i=1}^{\lfloor ns \rfloor} \log \varphi(a_i) + \sum_{j=1}^{\lfloor nt \rfloor} \log \psi(b_j) \right) \right] \\ &\quad - \frac{1}{n} \log \mathbb{P}_{\nu_1, \nu_2}(A_n). \end{aligned}$$

Note that for any measures ν_1, ν_2 , we have $g_{\nu_1, \nu_2}(s, t) > 0$, so we have not divided by zero above. The last term tends to zero because $\mathbb{P}_{\nu_1, \nu_2}(A_n) \rightarrow 1$ as $n \rightarrow \infty$. For the remaining term, we note that

$$\begin{aligned} &\mathbb{E}_{\nu_1, \nu_2} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n) \left(\sum_{i=1}^{\lfloor ns \rfloor} \log \varphi(a_i) + \sum_{j=1}^{\lfloor nt \rfloor} \log \psi(b_j) \right) \right] \\ &= \mathbb{E}_{\nu_1, \nu_2} \left[\left(\sum_{i=1}^{\lfloor ns \rfloor} \log \varphi(a_i) + \sum_{j=1}^{\lfloor nt \rfloor} \log \psi(b_j) \right) \right] \\ &\quad - \mathbb{E}_{\nu_1, \nu_2} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n^c) \log \left(\prod_{i=1}^{\lfloor ns \rfloor} \varphi(a_i) \prod_{j=1}^{\lfloor nt \rfloor} \psi(b_j) \right) \right] \\ &= \lfloor ns \rfloor \mathbb{H}(\nu_1 | \alpha) + \lfloor nt \rfloor \mathbb{H}(\nu_2 | \beta) \\ &\quad - \mathbb{E}_{\alpha, \beta} \left[\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n^c) \prod_{i=1}^{\lfloor ns \rfloor} \varphi(a_i) \prod_{j=1}^{\lfloor nt \rfloor} \psi(b_j) \log \left(\prod_{i=1}^{\lfloor ns \rfloor} \varphi(a_i) \prod_{j=1}^{\lfloor nt \rfloor} \psi(b_j) \right) \right]. \end{aligned}$$

But $x \log x \geq -1/e$ and $\mathbf{P}_{\mathbf{a}, \mathbf{b}}(A_n^c) \in [0, 1]$ so the last term is bounded above by a constant. Dividing by n and taking $\limsup_{n \rightarrow \infty}$, then optimizing over ν_1, ν_2 gives the result. \square

To show that the annealed model has non-trivial rate n large deviations to the left of the shape function, it suffices to show that there exists $\nu_1 \in \mathcal{M}^\alpha$ with $g_{\nu_1, \beta}(s, t) < g_{\alpha, \beta}(s, t)$. The next lemma gives mild conditions under which this is the case.

Lemma 5.1. *Suppose that α is not degenerate and $\mathbb{E}^\alpha[a \log a] < \infty$. Then there exists ν_1 with $\mathbb{H}(\nu_1 | \alpha) < \infty$ and $g_{\nu_1, \beta}(s, t) < g_{\alpha, \beta}(s, t)$.*

Proof. Define ν_1 by $(d\nu_1/d\alpha)(a) \simeq a$. Note that $\mathbb{H}(\nu_1 | \alpha) < \infty$ by hypothesis. Let $\zeta \in [-\alpha, \beta]$ be such that $g_{\alpha, \beta}(s, t) = s \mathbb{E}[(a + \zeta)^{-1}] + t \mathbb{E}[(b - \zeta)^{-1}]$. Because $\alpha \neq \delta_c$ for any c , the Cauchy–Schwarz inequality gives

$$1 = \mathbb{E} \left[\sqrt{a + \zeta} \sqrt{a + \zeta}^{-1} \right]^2 < \mathbb{E}[a + \zeta] \mathbb{E}[(a + \zeta)^{-1}].$$

Rearranging implies that $\mathbb{E}[a(a + \zeta)^{-1}] < \mathbb{E}[a] \mathbb{E}[(a + \zeta)^{-1}]$. It then follows that

$$g_{\nu_1, \beta}(s, t) \leq s \mathbb{E}[a]^{-1} \mathbb{E} \left[\frac{a}{a + \zeta} \right] + t \mathbb{E} \left[\frac{1}{b - \zeta} \right]$$

$$< s \mathbb{E} \left[\frac{1}{a + \zeta} \right] + t \mathbb{E} \left[\frac{1}{b - \zeta} \right] = g_{\alpha, \beta}(s, t).$$

□

We expect that the moment condition in the previous lemma is unnecessary.

6. Large deviation principle

We prove Theorem 2.3 by working with Legendre–Fenchel transforms and appealing to convex duality.

Lemma 6.1. *For all $s, t > 0$,*

$$\mathbf{J}_{s,t}^*(\lambda) = \begin{cases} \mathbf{L}_{s,t}(\lambda) & \lambda \geq 0, \\ \infty & \lambda < 0, \end{cases} \quad \mathbb{J}_{s,t}^*(\lambda) = \begin{cases} \mathbb{L}_{s,t}(\lambda) & \lambda \geq 0, \\ \infty & \lambda < 0. \end{cases}$$

Proof. We give the proof of the result under $\mathbf{P}_{\mathbf{a}, \mathbf{b}}$. The proof under \mathbb{P} is similar. Recall the regularity properties of $\mathbf{J}_{s,t}(\cdot)$ proven in Proposition A.1 in the appendix. The result for $\lambda < 0$ follows from the observation that $\mathbf{J}_{s,t}(r) = 0$ for $r \leq g(s, t)$. For all $\lambda > 0$, by the exponential Markov inequality we have

$$\frac{1}{n} \log \mathbf{P}_{\mathbf{a}, \mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) \leq \frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] - \lambda r.$$

Sending $n \rightarrow \infty$ gives $\lambda r - \mathbf{J}_{s,t}(r) \leq \mathbf{L}_{s,t}(\lambda)$ and taking $\sup_{r \in \mathbb{R}}$ implies $\mathbf{J}_{s,t}^*(\lambda) \leq \mathbf{L}_{s,t}(\lambda)$. For the reverse inequality, we next consider the case $\lambda \in (0, \underline{\alpha} + \underline{\beta})$. Fix $M > 0$ and let $\{x_i\}_{i=0}^K$ be a partition of $[0, M]$. We observe that

$$\begin{aligned} \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] &= \sum_{i=1}^K \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)} \mathbf{1}_{(x_{i-1}, x_i]}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor))] \\ &\quad + \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)} \mathbf{1}_{(M, \infty)}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor))]. \end{aligned}$$

Consequently, we see that

$$\begin{aligned} &\frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] \\ &\leq \max \left\{ \max_{0 \leq i \leq K} \left\{ \lambda x_i + \frac{1}{n} \log \mathbf{P}_{\mathbf{a}, \mathbf{b}}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq x_{i-1}) \right\}, \right. \\ &\quad \left. \frac{1}{n} \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)} \mathbf{1}_{(M, \infty)}(n^{-1}G(\lfloor ns \rfloor, \lfloor nt \rfloor))] \right\} \\ &\quad + \frac{K+1}{n}. \end{aligned}$$

Take $\limsup_{n \rightarrow \infty}$ then $K \rightarrow \infty$. Using continuity of $r \mapsto \mathbf{J}_{s,t}(r)$, we see that

$$\begin{aligned} \mathbf{L}_{s,t}(\lambda) &\leq \max_{0 \leq r \leq M} \{\lambda r - \mathbf{J}_{s,t}(r)\} \\ &\vee \limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}} [e^{\lambda G([ns],[nt])} \mathbf{1}_{(M,\infty)}(n^{-1}G([ns],[nt]))}]. \end{aligned}$$

Let $p, q > 1$ be such that $p^{-1} + q^{-1} = 1$ and $p\lambda < \underline{\alpha} + \underline{\beta}$. Then

$$\begin{aligned} &\frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}} [e^{\lambda G([ns],[nt])} \mathbf{1}_{(M,\infty)}(n^{-1}G([ns],[nt]))}] \\ &\leq \frac{1}{pn} \log \mathbf{E}_{\mathbf{a},\mathbf{b}} [e^{\lambda p G([ns],[nt])}] + \frac{1}{qn} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(n^{-1}G([ns],[nt]) \geq M). \end{aligned}$$

From this, we see that there exist deterministic constants C_1, C_2 such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a},\mathbf{b}} [e^{\lambda G([ns],[nt])} \mathbf{1}_{(M,\infty)}(n^{-1}G([ns],[nt]))}] \leq C_1 - C_2 \mathbf{J}_{s,t}(M).$$

Recall that $\lambda r \leq \mathbf{L}_{s,t}(\lambda) + \mathbf{J}_{s,t}(r)$, so that as $M \rightarrow \infty$, $\mathbf{J}_{s,t}(M) \rightarrow \infty$. Since $\max_{r \leq M} \{\lambda r - \mathbf{J}_{s,t}(r)\} \leq \mathbf{J}_{s,t}^*(\lambda)$, it follows that we have $\mathbf{L}_{s,t}(\lambda) \leq \mathbf{J}_{s,t}^*(\lambda)$.

Next, we turn to the case $\lambda = \underline{\alpha} + \underline{\beta}$. We observe that as $\lambda \uparrow \underline{\alpha} + \underline{\beta}$, $\mathbf{L}_{s,t}(\lambda) \uparrow \mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta})$. Suppose that $\mathbf{L}_{s,t}(r) < \infty$. Fix $\varepsilon > 0$ and take $\lambda < \underline{\alpha} + \underline{\beta}$ such that $\sup_{r \in \mathbb{R}} \{\lambda r - \mathbf{J}_{s,t}(r)\} = \mathbf{L}_{s,t}(\lambda) \geq \mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta}) - 2\varepsilon$. Then there exists $r > 0$ so that $\lambda r - \mathbf{J}_{s,t}(r) \geq \mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta}) - \varepsilon$. Since $(\underline{\alpha} + \underline{\beta})r > \lambda r$, it follows that $\mathbf{J}_{s,t}^*(\underline{\alpha} + \underline{\beta}) \geq \mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta}) - \varepsilon$. The case $\mathbf{L}_{s,t}(\underline{\alpha} + \underline{\beta}) = \infty$ is similar.

Finally, we consider the case $\lambda > \underline{\alpha} + \underline{\beta}$, where $\mathbf{L}_{s,t}(\lambda) = \infty$. For each (i, j) , we eventually have $G([ns], [nt]) \geq W(i, j)$. This implies that for all (i, j) , $\mathbf{J}_{s,t}(r) \leq (a_i + b_j)r \mathbf{1}_{\{r \geq 0\}}$ and therefore μ almost surely, $\mathbf{J}_{s,t}(r) \leq (\underline{\alpha} + \underline{\beta})r \mathbf{1}_{\{r \geq 0\}}$. Taking Legendre–Fenchel transforms of this inequality shows that $\mathbf{J}_{s,t}^*(\lambda) = \infty$. \square

Proof of Theorem 2.3. Proposition A.1 shows that $r \mapsto \mathbf{J}_{s,t}(r)$ and $r \mapsto \mathbb{J}_{s,t}(r)$ are real valued convex functions on \mathbb{R} . The result follows from taking Legendre–Fenchel transforms of the expressions in the previous lemma [20, Theorem 12.2]. \square

Proof of Theorem 2.4. Fix an open set $O \subset \mathbb{R}$

1. If $O \subset (-\infty, g(s, t))$ then there is nothing to prove by Lemma 2.1.
2. If $g(s, t) \in O$, then

$$\limsup_{n \rightarrow \infty} -n^{-1} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(n^{-1}G([ns],[nt]) \in O) = 0 = \inf_{r \in O} I_{s,t}(r)$$

3. If $O \cap (g(s, t), \infty) \neq \emptyset$, then $O \cap (g(s, t), \infty)$ contains an interval (r_0, r_1) . Note that

$$\mathbf{P}_{\mathbf{a},\mathbf{b}}(n^{-1}G([ns],[nt]) \in O)$$

$$\begin{aligned} &\geq \mathbf{P}_{\mathbf{a},\mathbf{b}}(G([ns], [nt]) \in (r_0, r_1)) \\ &= \mathbf{P}_{\mathbf{a},\mathbf{b}}(G([ns], [nt]) \geq r_0) - \mathbf{P}_{\mathbf{a},\mathbf{b}}(G([ns], [nt]) \geq r_1). \end{aligned}$$

Lemma 4.4 shows that $\mathbf{J}_{s,t}(r)$ is strictly increasing for $r > g(s, t)$, which implies that

$$\limsup_{n \rightarrow \infty} -n^{-1} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(n^{-1}G([ns], [nt]) \in O) \leq \mathbf{J}_{s,t}(r_0).$$

Let $r_n \in O \cap (g(s, t), \infty)$ be a sequence with $r_n \downarrow r_\infty = \inf\{x : x \in O \cap (g(s, t), \infty)\}$. Then because $\mathbf{J}_{s,t}(r)$ is continuous and non-decreasing, we see that

$$\begin{aligned} &\limsup_{n \rightarrow \infty} -n^{-1} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(n^{-1}G([ns], [nt]) \in O) \\ &\leq \mathbf{J}_{s,t}(r_\infty) = \inf_{r \in O \cap (g(s,t), \infty)} \mathbf{I}_{s,t}(r) = \inf_{r \in O} \mathbf{I}_{s,t}(r). \end{aligned}$$

The upper bound follows from the regularity of $\mathbf{J}_{s,t}$, Theorem 2.3 and Lemma 2.1. □

7. Relative entropy and the rate functions

We now turn to the proof of Theorem 2.5. Our argument proving this result is purely convex analytic and does not show the probabilistic interpretation mentioned before the statement of the theorem. We begin with a technical lemma.

Lemma 7.1. *For $r > 0$, the map $(\alpha, \beta) \mapsto \mathbf{I}_{s,t}^{\alpha,\beta}(r)$ is convex on $\mathcal{M}_1(\mathbb{R}_+)^2$.*

Proof. Using (2.3), one can check that $(\alpha, \beta) \mapsto g_{\alpha,\beta}(s, t)$ is concave on $\mathcal{M}(\mathbb{R}_+)^2$. Thus, $\{(\alpha, \beta) : g_{\alpha,\beta}(s, t) \geq r\}$ is convex. Define for $(\alpha, \beta) \in \mathcal{M}_1(\mathbb{R}_+)^2$

$$F(\alpha, \beta) = \sup_{\substack{\lambda \in (0, \alpha + \beta) \\ z \in (-\alpha, \beta - \lambda)}} \left\{ \lambda r - s \mathbf{E}^\alpha \left[\log \left(\frac{a + z + \lambda}{a + z} \right) \right] - t \mathbf{E}^\beta \left[\log \left(\frac{b - z}{b - z - \lambda} \right) \right] \right\}.$$

Fix $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathcal{M}_1(\mathbb{R}_+)$ and $\delta \in (0, 1)$. Denote by $\alpha^\delta = \delta\alpha_1 + (1 - \delta)\alpha_2$ and by $\beta^\delta = \delta\beta_1 + (1 - \delta)\beta_2$. Note that $\underline{\alpha}^\delta = \alpha_1 \wedge \alpha_2$ and $\underline{\beta}^\delta = \beta_1 \wedge \beta_2$. Then

$$\begin{aligned} &F(\alpha^\delta, \beta^\delta) \\ &= \sup_{\substack{\lambda \in (0, \alpha^\delta + \beta^\delta) \\ z \in (-\underline{\alpha}^\delta, \underline{\beta}^\delta - \lambda)}} \left\{ \lambda r - s \mathbf{E}^{\alpha^\delta} \left[\log \left(\frac{a + z + \lambda}{a + z} \right) \right] - t \mathbf{E}^{\beta^\delta} \left[\log \left(\frac{b - z}{b - z - \lambda} \right) \right] \right\} \end{aligned}$$

$$\begin{aligned}
&\leq \delta \sup_{\substack{\lambda \in (0, \underline{\alpha}^\delta + \underline{\beta}^\delta) \\ z \in (-\underline{\alpha}^\delta, \underline{\beta}^\delta - \lambda)}} \left\{ \lambda r - s \mathbf{E}^{\alpha_1} \left[\log \left(\frac{a+z+\lambda}{a+z} \right) \right] - t \mathbf{E}^{\beta_1} \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] \right\} \\
&\quad + (1-\delta) \sup_{\substack{\lambda \in (0, \underline{\alpha}^\delta + \underline{\beta}^\delta) \\ z \in (-\underline{\alpha}^\delta, \underline{\beta}^\delta - \lambda)}} \left\{ \lambda r - s \mathbf{E}^{\alpha_2} \left[\log \left(\frac{a+z+\lambda}{a+z} \right) \right] - t \mathbf{E}^{\beta_2} \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] \right\} \\
&\leq \delta \sup_{\substack{\lambda \in (0, \underline{\alpha}_1 + \underline{\beta}_1) \\ z \in (-\underline{\alpha}_1, \underline{\beta}_1 - \lambda)}} \left\{ \lambda r - s \mathbf{E}^{\alpha_1} \left[\log \left(\frac{a+z+\lambda}{a+z} \right) \right] - t \mathbf{E}^{\beta_1} \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] \right\} \\
&\quad + (1-\delta) \sup_{\substack{\lambda \in (0, \underline{\alpha}_2 + \underline{\beta}_2) \\ z \in (-\underline{\alpha}_2, \underline{\beta}_2 - \lambda)}} \left\{ \lambda r - s \mathbf{E}^{\alpha_2} \left[\log \left(\frac{a+z+\lambda}{a+z} \right) \right] - t \mathbf{E}^{\beta_2} \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] \right\},
\end{aligned}$$

so that F is convex on $\mathcal{M}(\mathbb{R}^+)^2$. Then we see from (2.11) that $(\alpha, \beta) \mapsto \mathbf{I}_{s,t}^{\alpha,\beta}(r)$ is convex on $\mathcal{M}(\mathbb{R}^+)^2$. \square

Proof of Theorem 2.5. Theorem 2.3 and the variational characterization of relative entropy, [19, Theorem 5.4], imply that for $r > g(s, t)$,

$$\begin{aligned}
\mathbb{J}_{s,t}^{\alpha,\beta}(r) &= \sup_{\substack{\lambda \in (0, \underline{\alpha} + \underline{\beta}) \\ z \in (-\underline{\alpha}, \underline{\beta} - \lambda)}} \left\{ \lambda r - s \log \mathbf{E}^\alpha \left[\frac{a+z+\lambda}{a+z} \right] - t \log \mathbf{E}^\beta \left[\frac{b-z}{b-z-\lambda} \right] \right\} \\
&= \sup_{\substack{\lambda \in (0, \underline{\alpha} + \underline{\beta}) \\ z \in (-\underline{\alpha}, \underline{\beta} - \lambda)}} \inf_{\substack{\nu_1 \in \mathcal{M}^\alpha \\ \nu_2 \in \mathcal{M}^\beta}} \left\{ \lambda r - s \mathbf{E}^{\nu_1} \left[\log \left(\frac{a+z+\lambda}{a+z} \right) \right] \right. \\
&\quad \left. - t \mathbf{E}^{\nu_2} \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] + s \mathbf{H}(\nu_1 | \alpha) + t \mathbf{H}(\nu_2 | \beta) \right\} \\
&\leq \inf_{\substack{\nu_1 \in \mathcal{M}^\alpha \\ \nu_2 \in \mathcal{M}^\beta}} \sup_{\substack{\lambda \in (0, \underline{\alpha} + \underline{\beta}) \\ z \in (-\underline{\alpha}, \underline{\beta} - \lambda)}} \left\{ \lambda r - s \mathbf{E}^{\nu_1} \left[\log \left(\frac{a+z+\lambda}{a+z} \right) \right] \right. \\
&\quad \left. - t \mathbf{E}^{\nu_2} \left[\log \left(\frac{b-z}{b-z-\lambda} \right) \right] + s \mathbf{H}(\nu_1 | \alpha) + t \mathbf{H}(\nu_2 | \beta) \right\}.
\end{aligned}$$

Note that if $\nu_1 \ll \alpha$, it must be the case that $\nu_1 \geq \underline{\alpha}$ and similarly, $\nu_2 \geq \underline{\beta}$. It follows that we may extend the region in the inner supremum to obtain

$$\mathbb{J}_{s,t}^{\alpha,\beta}(r) \leq \inf_{\nu_1, \nu_2} \left\{ \mathbf{I}_{s,t}^{\nu_1, \nu_2}(r) + s \mathbf{H}(\nu_1 | \alpha) + t \mathbf{H}(\nu_2 | \beta) \right\}.$$

The map $(\nu_1, \nu_2) \mapsto \mathbf{I}_{s,t}^{\nu_1, \nu_2}(r) + s \mathbf{H}(\nu_1 | \alpha) + t \mathbf{H}(\nu_2 | \beta)$ is strictly convex on the convex set $\mathcal{M}^\alpha \times \mathcal{M}^\beta$ so at most one minimizing pair (ν_1, ν_2) exists. It therefore

suffices to show that we have equality with the measures ν_1, ν_2 defined in the statement of the theorem. We argue this by cases.

A maximizing pair λ_\star, z_\star satisfying $\lambda_\star \in [0, \underline{\alpha} + \underline{\beta}]$, $z_\star \in [-\underline{\alpha}, \underline{\beta} - \lambda_\star]$ exist for the annealed right-tail rate function by Lemmas 4.2 and 4.4. (z_\star denotes $z_\star(\lambda_\star)$ in the notation of Section 4. Also, by Corollary 3.1, $z_\star(\underline{\alpha} + \underline{\beta}) = -\underline{\alpha}$.) Note that $\lambda_\star = 0$ is impossible because $\mathbb{J}_{s,t}^{\alpha,\beta}(r) > 0$ by Lemma 4.4. If $\lambda_\star \in (0, \underline{\alpha} + \underline{\beta})$ and $z_\star \in (-\underline{\alpha}, \underline{\beta} - \lambda_\star)$, then $\nu_1 \in \mathcal{M}^\alpha$ and $\nu_2 \in \mathcal{M}^\beta$ because their densities with respect to α and β are bounded. Taking derivatives in (2.12), we see that z_\star and λ_\star solve

$$0 = s \mathbb{E}^{\nu_1} \left[\frac{1}{a + z_\star} - \frac{1}{a + z_\star + \lambda_\star} \right] + t \mathbb{E}^{\nu_2} \left[\frac{1}{b - z_\star} - \frac{1}{b - z_\star - \lambda_\star} \right], \quad (7.1)$$

$$0 = r - s \mathbb{E}^{\nu_1} \left[\frac{1}{a + z_\star + \lambda_\star} \right] - t \mathbb{E}^{\nu_2} \left[\frac{1}{b - z_\star - \lambda_\star} \right]. \quad (7.2)$$

These are precisely the first order conditions implying that

$$\mathbf{I}_{s,t}^{\nu_1, \nu_2}(r) = \lambda_\star r - s \mathbb{E}^{\nu_1} \left[\log \frac{a + z_\star + \lambda_\star}{a + z_\star} \right] - t \mathbb{E}^{\nu_2} \left[\log \frac{b - z_\star}{b - z_\star - \lambda_\star} \right].$$

The definition of relative entropy and a little algebra then show that

$$\mathbb{J}_{s,t}^{\alpha,\beta}(r) = \mathbf{I}_{s,t}^{\nu_1, \nu_2}(r) + s \mathbb{H}(\nu_1 | \alpha) + t \mathbb{H}(\nu_2 | \alpha).$$

The remaining cases are similar in that once we know that the extremizers are the same for $\mathbb{J}_{s,t}^{\alpha,\beta}(r)$ and $\mathbf{I}_{s,t}^{\nu_1, \nu_2}(r)$, the result follows. The necessary and sufficient conditions in Lemmas 4.1 and 4.2 show that ν_1 and ν_2 are well defined and that this equality continues to hold if $\lambda_\star < \underline{\alpha} + \underline{\beta}$ and $z_\star = -\underline{\alpha}$ or $z_\star = \underline{\beta} - \lambda_\star$. The only remaining case is $\lambda_\star = \underline{\alpha} + \underline{\beta}$ and $z_\star = -\underline{\alpha}$. $\lambda_\star = \underline{\alpha} + \underline{\beta}$ is equivalent to $r \geq (\mathbb{L}_{s,t}^{\alpha,\beta})'(\underline{\alpha} + \underline{\beta})$. By Corollary 4.1, this condition implies that ν_1 and ν_2 are well defined and $(\mathbf{L}_{s,t}^{\nu_1, \nu_2})'(\underline{\alpha} + \underline{\beta}) = (\mathbb{L}_{s,t}^{\alpha,\beta})'(\underline{\alpha} + \underline{\beta})$. The result follows. \square

8. Scaling estimates

In this section, we prove the scaling estimates for the quenched and the annealed rate functions. See the discussion Section 4 for the notation below. If $c_1 < s/t < c_2$ we have $\partial_z g_\zeta(s, t) = 0$ and, therefore,

$$g_z(s, t) = g(s, t) + \partial_z^2 g_\zeta(s, t)(z - \zeta)^2/2 + o((z - \zeta)^2). \quad (8.1)$$

In fact, (8.1) holds for $s/t = c_1$ and $s/t = c_2$ as well provided that

$$\mathbb{E} \left[\frac{1}{(a - \underline{\alpha})^3} \right] < \infty, \quad \mathbb{E} \left[\frac{1}{(b - \underline{\beta})^3} \right] < \infty; \quad (8.2)$$

that is, assuming that $\partial_z^2 g_z(s, t)$ has limits at the endpoints $-\underline{\alpha}$ and $\underline{\beta}$.

Proof of Theorem 2.7. For $\varepsilon > 0$ sufficiently small, we have

$$\mathbf{I}'_{s,t}(r) = \lambda_*(r), \quad \mathbf{L}'_{s,t}(\lambda_*(r)) = r \tag{8.3}$$

whenever $g(s, t) \leq r \leq g(s, t) + \varepsilon$ by Lemma 4.4. We begin with the case $c_1 < s/t < c_2$. Then $\zeta \in (-\alpha, \beta)$. We recall λ_1 and λ_2 defined in Lemma 4.1. Because $\partial_z F(-\alpha, 0) = \partial_z g_{-\alpha}(s, t) < 0$ and $\partial_z F(-\alpha, 0) = \partial_z g_{-\alpha}(s, t) > 0$, we conclude that $\lambda_1 > 0$ and $\lambda_2 > 0$. Hence,

$$\begin{aligned} z_*'(\lambda) &= -\frac{\partial_\lambda \partial_z F(z_*, \lambda)}{\partial_z^2 F(z_*, \lambda)} \\ &= -\frac{s \mathbb{E}\left[\frac{1}{(a+z_*)(a+z_*+\lambda)^2}\right] + t \mathbb{E}\left[\frac{1}{(b-z_*)(b-z_*-\lambda)^2}\right]}{s \mathbb{E}\left[\frac{2a+2z_*+\lambda}{(a+z_*+\lambda)^2(a+z_*)^2}\right] + t \mathbb{E}\left[\frac{2b-2z_*-\lambda}{(b-z_*-\lambda)^2(b-z_*)^2}\right]} \end{aligned}$$

for $0 < \lambda < \lambda_1 \wedge \lambda_2$. Letting $\lambda \downarrow 0$ yields $z_*'(0^+) = -1/2$. It follows that $z_*(\lambda) = \zeta - \lambda/2 + o(\lambda)$ as $\lambda \downarrow 0$. We obtain $\mathbf{L}'_{s,t}(\lambda) = g_{z_*+\lambda}(s, t) = g(s, t) + \partial_z^2 g_\zeta(s, t)\lambda^2/8 + o(\lambda^2)$ as $\lambda \downarrow 0$. Then,

$$\mathbf{I}'_{s,t}(g(s, t) + \varepsilon) = \frac{2\sqrt{2}}{\sqrt{\partial_z^2 g_\zeta(s, t)}} \varepsilon^{1/2} + o(\varepsilon^{1/2}),$$

and integrating gives

$$\begin{aligned} \mathbf{I}_{s,t}(g(s, t) + \varepsilon) &= \frac{4\sqrt{2}\varepsilon^{3/2}}{3\sqrt{\partial_z^2 g_\zeta(s, t)}} + o(\varepsilon^{3/2}) \\ &= \frac{4}{3} \frac{\varepsilon^{3/2}}{\sqrt{s \mathbb{E}\left[\frac{1}{(a+\zeta)^3}\right] + t \mathbb{E}\left[\frac{1}{(b-\zeta)^3}\right]}} + o(\varepsilon^{3/2}) \end{aligned} \tag{8.4}$$

as $\varepsilon \downarrow 0$. Now, suppose that $s/t \leq c_1$. Then $E[(a - \alpha)^{-2}] < \infty$, $\zeta = -\alpha$ and $z_* = -\alpha$. Under condition (8.2), when $c_1 = s/t$, $\mathbf{L}'_{s,t}(\lambda) = g_{-\alpha+\lambda}(s, t) = g(s, t) + \partial_z^2 g_{-\alpha}(s, t)\lambda^2/2 + o(\lambda^2)$ and we reach (8.4) multiplied with 1/2. If $c_1 > s/t$ then $\partial_z g_{-\alpha}(s, t) > 0$ and we have $\mathbf{L}'_{s,t}(\lambda) = g_{-\alpha+\lambda}(s, t) = g(s, t) + \partial_z g_{-\alpha}(s, t)\lambda + o(\lambda)$. This leads to

$$\begin{aligned} \mathbf{I}_{s,t}(g(s, t) + \varepsilon) &= \frac{\varepsilon^2}{2\partial_z g_{-\alpha}(s, t)} + o(\varepsilon^2) \\ &= \frac{1}{2} \frac{\varepsilon^2}{-s \mathbb{E}\left[\frac{1}{(a-\alpha)^2}\right] + t \mathbb{E}\left[\frac{1}{(b+\alpha)^2}\right]} + o(\varepsilon^2). \end{aligned}$$

Analysis of the case $s/t \geq c_2$ is similar. □

Proof of Theorem 2.8. In the case $c_1 < s/t < c_2$, Hölder's inequality gives

$$\begin{aligned} \lim_{\lambda \downarrow 0} z_\star'(\lambda) &= - \lim_{\lambda \downarrow 0} \frac{\partial_\lambda \partial_z F(z_\star, \lambda)}{\partial_z^2 F(z_\star, \lambda)} \\ &= - \frac{s \mathbb{E}\left[\frac{1}{(a+\zeta)^2}\right] \mathbb{E}\left[\frac{1}{a+\zeta}\right] + 2t \mathbb{E}\left[\frac{1}{(b-\zeta)^3}\right] - t \mathbb{E}\left[\frac{1}{(b-\zeta)^2}\right] \mathbb{E}\left[\frac{1}{b-\zeta}\right]}{2s \mathbb{E}\left[\frac{1}{(a+\zeta)^3}\right] + 2t \mathbb{E}\left[\frac{1}{(b-\zeta)^3}\right]} \\ &\leq - \frac{s \mathbb{E}\left[\frac{1}{(a+\zeta)^2}\right] \mathbb{E}\left[\frac{1}{a+\zeta}\right] + t \mathbb{E}\left[\frac{1}{(b-\zeta)^3}\right]}{2s \mathbb{E}\left[\frac{1}{(a+\zeta)^3}\right] + 2t \mathbb{E}\left[\frac{1}{(b-\zeta)^3}\right]}. \end{aligned}$$

Hence, $z_\star(\lambda) = \zeta + c\lambda + o(\lambda)$, where $c < 0$. We have

$$\mathbb{J}'_{s,t}(\lambda) = s \frac{\mathbb{E}\left[\frac{1}{a+z_\star}\right]}{\mathbb{E}\left[\frac{a+z_\star+\lambda}{a+z_\star}\right]} + t \frac{\mathbb{E}\left[\frac{b-z_\star}{(b-z_\star-\lambda)^2}\right]}{\mathbb{E}\left[\frac{b-z_\star}{b-z_\star-\lambda}\right]} \quad (8.5)$$

$$= g_{z_\star+\lambda}(s, t) + \lambda \left(s \operatorname{Var}\left[\frac{1}{a+\zeta}\right] + t \operatorname{Var}\left[\frac{1}{b-\zeta}\right] \right) + o(\lambda) \quad (8.6)$$

$$= g(s, t) + \lambda \left(s \operatorname{Var}\left[\frac{1}{a+\zeta}\right] + t \operatorname{Var}\left[\frac{1}{b-\zeta}\right] \right) + o(\lambda). \quad (8.7)$$

Then, arguing as in the preceding proof, we obtain

$$\mathbb{J}_{s,t}(g(s, t) + \varepsilon) = \frac{1}{2} \frac{\varepsilon^2}{s \operatorname{Var}\left[\frac{1}{a+\zeta}\right] + t \operatorname{Var}\left[\frac{1}{b-\zeta}\right]} + o(\varepsilon). \quad (8.8)$$

Now consider $s/t \leq c_1$. Then $\zeta = z_\star = -\alpha$ and (8.5) still holds. If $s/t = c_1$ subsequent arguments go through assuming (8.2). This condition is needed in step (8.6), which relies on (8.1) with $\zeta = -\alpha$. Hence, we have (8.8). If $s/t < c_1$ then the coefficient of λ in (8.7) has an additional term $\partial_z g_{-\alpha}(s, t) > 0$, which leads to

$$\mathbb{J}_{s,t}(g(s, t) + \varepsilon) = \frac{1}{2} \frac{\varepsilon^2}{-s \mathbb{E}\left[\frac{1}{a-\alpha}\right]^2 + t \operatorname{Var}\left[\frac{1}{b+\alpha}\right] + t \mathbb{E}\left[\frac{1}{(b+\alpha)^2}\right]} + o(\varepsilon).$$

The case $s/t \geq c_2$ is analyzed similarly. \square

A. Right tail rate functions and Lyapunov exponents

Proposition A.1.

(a) μ -a.s., for $s, t > 0$ and $r \in \mathbb{R}$, there exists (nonrandom) $\mathbf{J}_{s,t}(r) \in [0, \infty)$ such that

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) = \mathbf{J}_{s,t}(r). \quad (\text{A.1})$$

(b) For all $s, t > 0$ and $r \in \mathbb{R}$, there exists $\mathbb{J}_{s,t}(r) \in [0, \infty)$ such that

$$\lim_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{P}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) = \mathbb{J}_{s,t}(r). \quad (\text{A.2})$$

(c) \mathbf{J} and \mathbb{J} are convex and homogeneous in (s, t, r) , nonincreasing in (s, t) and nondecreasing in r .

Proof. Fix $r \in \mathbb{R}$ and $s, t \in \mathbb{N}$. For integers $0 \leq m < n$, define

$$X_{m,n} = -\log \mathbf{P}_{\tau_{ms}(\mathbf{a}), \tau_{mt}(\mathbf{b})}(G((n-m)s, (n-m)t) \geq (n-m)r).$$

We verify that $\{X_{m,n}\}$ satisfy the hypotheses of the subadditive ergodic theorem in [17]. For subadditivity, note that

$$\begin{aligned} X_{0,n} &= -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(ns, nt) \geq nr) \\ &\leq -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(ms, mt) \geq mr) \\ &\quad -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G((n-m)s, (n-m)t) \circ \theta_{ms,mt} \geq (n-m)r) \\ &= X_{0,m} + X_{m,n}. \end{aligned}$$

For $k \in \mathbb{N}$, by the ergodicity assumptions on μ , the sequence $(X_{k,k+n})_{n \in \mathbb{N}}$ has the same distribution as $(X_{0,n})_{n \in \mathbb{N}}$ and the sequence $(X_{(n-1)k, nk})_{n \in \mathbb{N}}$ is ergodic. Moreover, $X_{0,n} \geq 0$ and

$$\mathbb{E} X_{0,n} \leq \mathbb{E} [-\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(W(1, 1) \geq nr)] = n \max\{r, 0\} \mathbb{E}[a + b] < \infty. \quad (\text{A.3})$$

Hence, by the subadditive ergodic theorem, (A.1) holds μ -a.s. (and in expectation under μ) with

$$\mathbf{J}_{s,t}(r) = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} X_{0,n} = \lim_{n \rightarrow \infty} -\frac{1}{n} \mathbb{E} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(ns, nt) \geq nr). \quad (\text{A.4})$$

We record some properties of $\mathbf{J}_{s,t}(r)$ for $s, t \in \mathbb{N}$ and $r \in \mathbb{R}$. It is clear from (A.4) that $\mathbf{J}_{s,t}(r)$ is nonincreasing in (s, t) and nondecreasing in r . In addition, $\mathbf{J}_{s,t}(r) = 0$ for $r \leq 0$ as G is nonnegative, and $\mathbf{J}_{cs,ct}(cr) = c \mathbf{J}_{s,t}(r)$ for $c \in \mathbb{N}$. By (A.3), $\mathbf{J}_{s,t}(r) \leq r \mathbb{E}[a + b] < \infty$ for $r \geq 0$. Also, for $s_1, s_2, t_1, t_2 \in \mathbb{N}$ and $r_1, r_2 \in \mathbb{R}$, we have

$$\mathbb{E} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(n(s_1 + s_2), n(t_1 + t_2)) \geq n(r_1 + r_2))$$

$$\geq \mathbb{E} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(ns_1, nt_1) \geq nr_1) \cdot \mathbb{E} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(ns_2, nt_2) \geq nr_2)$$

for $n \in \mathbb{N}$, which gives $\mathbf{J}_{s_1+s_2, t_1+t_2}(r_1+r_2) \geq \mathbf{J}_{s_1, t_1}(r_1) + \mathbf{J}_{s_2, t_2}(r_2)$. Then, for $0 \leq r \leq r' \leq r + 1/n$,

$$\begin{aligned} \mathbf{J}_{s,t}(r') - \mathbf{J}_{s,t}(r) &\leq \mathbf{J}_{s,t}(r + 1/n) - \mathbf{J}_{s,t}(r) \\ &= \frac{1}{n+1} \mathbf{J}_{(n+1)s, (n+1)t}(nr + r + 1 + 1/n) - \mathbf{J}_{s,t}(r) \\ &\leq \frac{\mathbf{J}_{ns, nt}(nr)}{n+1} - \mathbf{J}_{s,t}(r) + \frac{\mathbf{J}_{s,t}(r+2)}{n+1} \\ &= \frac{\mathbf{J}_{s,t}(r+2) - \mathbf{J}_{s,t}(r)}{n+1} \leq \frac{2r+2}{n} \mathbb{E}[a+b], \end{aligned} \tag{A.5}$$

which shows continuity of $\mathbf{J}_{s,t}(r)$ in r .

There exists a μ -a.s. event E on which (A.1) holds for all $s, t \in \mathbb{N}$ and $r \in \mathbb{Q}$. It follows from the monotonicity of $\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(ns, nt) \geq nr)$ in r and continuity of $\mathbf{J}_{s,t}$ that (A.1) holds for all $s, t \in \mathbb{N}$ and $r \in \mathbb{R}$ on E . From now on, let us work with $(\mathbf{a}, \mathbf{b}) \in E$.

For $c > 0$, $\delta \in (0, 1)$ and large enough $n \in \mathbb{N}$, we have

$$\begin{aligned} -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ncs \rfloor, \lfloor nct \rfloor) \geq nr) &\leq -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor cn \rfloor s, \lfloor cn \rfloor t) \geq \lfloor cn \rfloor r(1+\delta)), \\ -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ncs \rfloor, \lfloor nct \rfloor) \geq nr) &\geq -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor cn \rfloor s, \lfloor cn \rfloor t) \geq \lfloor cn \rfloor r(1-\delta)). \end{aligned} \tag{A.6}$$

It follows from these inequalities and continuity of $\mathbf{J}_{s,t}$ that (A.1) holds on E with $\mathbf{J}_{cs, ct}(cr) = c\mathbf{J}_{s,t}(r)$. In particular, $\mathbf{J}_{s,t}(r)$ exists for rational $s, t > 0$. Moreover, by homogeneity, the properties of $\mathbf{J}_{s,t}(r)$ noted in preceding paragraph hold for rational $s, t > 0$ as well.

For $s, t, \delta > 0$, choose rational s', t' such that $s'/(1+\delta) < s \leq s'$ and $t'/(1+\delta) < t \leq t'$. Then

$$\begin{aligned} -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) &\geq -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns' \rfloor, \lfloor nt' \rfloor) \geq nr), \\ -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) &\leq -\log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns'/(1+\delta) \rfloor, \lfloor nt'/(1+\delta) \rfloor) \geq nr). \end{aligned} \tag{A.7}$$

It follows that

$$\begin{aligned} \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) &\geq \mathbf{J}_{s', t'}(r), \\ \limsup_{n \rightarrow \infty} -\frac{1}{n} \log \mathbf{P}_{\mathbf{a},\mathbf{b}}(G(\lfloor ns \rfloor, \lfloor nt \rfloor) \geq nr) &\leq \mathbf{J}_{s'/(1+\delta), t'/(1+\delta)}(r) \\ &= \mathbf{J}_{s', t'}((1+\delta)r)/(1+\delta). \end{aligned}$$

Using (A.5), we obtain

$$\frac{\mathbf{J}_{s', t'}((1+\delta)r)}{1+\delta} - \mathbf{J}_{s', t'}(r) \leq \mathbf{J}_{s', t'}((1+\delta)r) - \mathbf{J}_{s', t'}(r) \leq \frac{2r+2}{\lfloor (r\delta)^{-1} \rfloor} \mathbb{E}[a+b].$$

As $\delta \downarrow 0$, we have $s' \downarrow s$ and $t' \downarrow t$. Hence, we conclude that $\mathbf{J}_{s,t}(r)$ exists and equals the limit of $\mathbf{J}_{s',t'}(r)$, and also enjoys the properties of mentioned above. Finally, it follows from subadditivity and homogeneity that \mathbf{J} is convex. \square

Proposition A.2.

(a) μ -a.s., for any $s, t > 0$ and $\lambda \in \mathbb{R}$, there exists $\mathbf{L}_{s,t}(\lambda) \in [-\infty, \infty]$ such that,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] = \mathbf{L}_{s,t}(\lambda). \quad (\text{A.8})$$

(b) For any $s, t > 0$ and $\lambda \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{E} [e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] = \mathbb{L}_{s,t}(\lambda). \quad (\text{A.9})$$

(c) $\mathbf{L}_{s,t}(\lambda)$ and $\mathbb{L}_{s,t}(\lambda)$ are nondecreasing and convex in λ .

(d) $\lambda \mathbf{L}_{s,t}(\lambda)$ and $\lambda \mathbb{L}_{s,t}(\lambda)$ are nondecreasing, homogeneous and concave in (s, t) .

Proof. Fix $\lambda \in \mathbb{R}$ and $s, t \in \mathbb{N}$. Define

$$X_{m,n} = -\lambda \log \mathbf{E}_{\tau_{ms}(\mathbf{a}), \tau_{mt}(\mathbf{b})} [e^{\lambda G((n-m)s, (n-m)t)}]$$

for integers $0 \leq m < n$. Then $\{X_{m,n} : 0 \leq m < n\}$ are nonpositive and subadditive, and the conditions of the subadditive ergodic theorem are in place to claim the existence of $\mathbf{L}_{s,t}(\lambda) \in [-\infty, \infty]$ such that (A.8) holds μ -a.s.

For $\lambda \in \mathbb{R}$, $s, t \in \mathbb{N}$ and $c > 0$, we have

$$\begin{aligned} -\lambda \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor nc \rfloor s, \lfloor nc \rfloor t)}] &\leq -\lambda \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor ncs \rfloor, \lfloor nct \rfloor)}] \\ &\leq -\lambda \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor nc \rfloor s, \lfloor nc \rfloor t)}]. \end{aligned}$$

Also, for $\lambda \in \mathbb{R}$, $s, s', t, t', \delta > 0$ such that s', t' are rational, $s'/(1+\delta) < s \leq s'$ and $t'/(1+\delta) < t \leq t'$,

$$\begin{aligned} -\lambda \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor ns' \rfloor, \lfloor nt' \rfloor)}] &\leq -\lambda \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{\lambda G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] \\ &\leq -\lambda \log \mathbf{E}_{\mathbf{a}, \mathbf{b}} \left[\exp \left\{ \lambda G \left(\left\lfloor \frac{ns'}{1+\delta} \right\rfloor, \left\lfloor \frac{nt'}{1+\delta} \right\rfloor \right) \right\} \right]. \end{aligned}$$

Using these inequalities as in the preceding proof, we obtain (A.8) for all $s, t > 0$ μ -a.s. and the claimed properties of the function $(s, t) \mapsto \lambda \mathbf{L}_{s,t}(\lambda)$.

Now fix $s, t > 0$. Note that $\mathbf{L}_{s,t}(\lambda)$ is nondecreasing in λ . Let $\lambda_0 = \sup_{\lambda \in \mathbb{R}} \{\mathbf{L}_{s,t}(\lambda) < \infty\}$. For $\lambda_1, \lambda_2 \in \mathbb{R}$ and $c_1, c_2 \in (0, 1)$ with $c_1 + c_2 = 1$, by Hölder's inequality,

$$\log \mathbf{E}_{\mathbf{a}, \mathbf{b}} [e^{(c_1 \lambda_1 + c_2 \lambda_2) G(\lfloor ns \rfloor, \lfloor nt \rfloor)}]$$

$$\leq c_1 \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda_1 G(\lfloor ns \rfloor, \lfloor nt \rfloor)}] + c_2 \log \mathbf{E}_{\mathbf{a}, \mathbf{b}}[e^{\lambda_2 G(\lfloor ns \rfloor, \lfloor nt \rfloor)}],$$

which implies that $\mathbf{L}_{s,t}(c_1 \lambda_1 + c_2 \lambda_2) \leq c_1 \mathbf{L}_{s,t}(\lambda_1) + c_2 \mathbf{L}_{s,t}(\lambda_2)$. Hence, $\mathbf{L}_{s,t}(\lambda)$ is continuous in λ on $(-\infty, \lambda_0)$. Using this and the monotonicity of last-passage times, we deduce that (A.8) holds for all $s, t > 0$ and $\lambda \in \mathbb{R}$ μ -a.s. \square

Acknowledgements

The authors are grateful to Timo Seppäläinen for many helpful conversations. The authors would also like to thank the Department of Mathematics at the University of Wisconsin–Madison, where much of this work was completed.

References

- [1] G.W. ANDERSON, A. GUIONNET AND O. ZEITOUNI (2010) *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics **118**. Cambridge University Press, Cambridge.
- [2] A. BORODIN, I. CORWIN AND P. FERRARI (2014) Free energy fluctuations for directed polymers in random media in 1 + 1 dimension. *Comm. Pure Appl. Math.* **67** (7), 1129–1214.
- [3] H. COHN, N. ELKIES AND J. PROPP (1996) Local statistics for random domino tilings of the Aztec diamond. *Duke Math. J.* **85** (1), 117–166.
- [4] F. COMETS, N. GANTERT AND O. ZEITOUNI (2000) Quenched, annealed and functional large deviations for one-dimensional random walk in random environment. *Probab. Theory Relat. Fields* **118**, 65–114.
- [5] I. CORWIN (2011) The Kardar–Parisi–Zhang equation and universality class. *Random Matrices: Theory and Applications* **1**.
- [6] I. CORWIN, T. SEPPÄLÄINEN AND H. SHEN (2015) The strict-weak lattice polymer. *J. Stat. Phys.* **160** (4), 1027–1053.
- [7] A. DEMBO AND O. ZEITOUNI (1998) *Large Deviations Techniques and Applications*. 2nd edn. Springer.
- [8] E. EMRAH (2016) Limit shapes for inhomogeneous corner growth models with exponential and geometric weights. *Electron. Comm. Probab.* **21** (42), 1–16.
- [9] N. GEORGIU AND T. SEPPÄLÄINEN (2013) Large deviation rate functions for the partition function in a log-gamma distributed random potential. *Ann. Probab.* **41** (6), 4248–4286.
- [10] J. GRAVNER, C.A. TRACY AND H. WIDOM (2001) Limit theorems for height fluctuations in a class of discrete space and time growth models. *J. Statist. Phys.* **102** (5–6), 1085–1132.
- [11] J. GRAVNER, C.A. TRACY AND H. WIDOM (2002) Fluctuations in the composite regime of a disordered growth model. *Comm. Math. Phys.* **229** (3), 433–458.

- [12] J. GRAVNER, C.A. TRACY AND H. WIDOM (2002) A growth model in a random environment. *Ann. Probab.* **30** (3), 1340–1368.
- [13] C. JANJIGIAN (2015) Large deviations of the free energy in the O’Connell–Yor polymer. *J. Stat. Phys.* **160** (4), 1054–1080.
- [14] K. JOHANSSON (2000) Shape fluctuations and random matrices. *Commun. Math. Phys.*, 437–476.
- [15] W. JOSCKUSCH, J. PROPP AND P. SHOR (1998) Random domino tilings and the arctic circle theorem. Preprint [arXiv:math/9801068](https://arxiv.org/abs/math/9801068).
- [16] H. KESTEN (1986) Aspects of first passage percolation. In: *École d’été de probabilités de Saint-Flour, XIV—1984, Lect. Notes Math.* **1180**, 125–264.
- [17] T.M. LIGGETT (1985) An improved subadditive ergodic theorem. *Ann. Probab.* **13** (4), 1279–1285.
- [18] J.B. MARTIN (2005). Last-passage percolation with general weight distribution. www.stats.ox.ac.uk/~martin/papers/lppsurvey2.ps.
- [19] F. RASSOUL-AGHA AND T. SEPPÄLÄINEN (2015) *A Course on Large Deviations with an Introduction to Gibbs Measures*. Graduate Studies in Mathematics **162**. American Mathematics Society.
- [20] R.T. ROCKAFELLAR (1970) *Convex Analysis*, Princeton Mathematical Series **28**. Princeton University Press.
- [21] H. ROST (1981) Nonequilibrium behavior of a many particle process: Density profile and local equilibria. *Z. Wahrsch. Verw. Gebiete* **58**, 41–53.
- [22] T. SEPPÄLÄINEN (1998) Coupling the totally asymmetric simple exclusion process with a moving interface. *Markov Process. Relat. Fields* **4** (4), 593–628.
- [23] T. SEPPÄLÄINEN (1998) Hydrodynamic scaling, convex duality and asymptotic shapes of growth models. *Markov Process. Relat. Fields* **4** (1), 1–26.
- [24] T. SEPPÄLÄINEN (1998) Large deviations for increasing sequences on the plane. *Probab. Theory Relat. Fields* **112** (2), 221–244.
- [25] T. SEPPÄLÄINEN (2009). *Lecture Notes on the Corner Growth Model*. <http://www.math.wisc.edu/~seppalai/cornergrowth-book/ajo.pdf>.
- [26] T. SEPPÄLÄINEN (2012) Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* **40** (1), 19–73.
- [27] T. SEPPÄLÄINEN AND J. KRUG (1999) Hydrodynamics and platoon formation for a totally asymmetric exclusion model with particlewise disorder. *J. Stat. Phys.* **95**, 525–567.
- [28] M. SION (1958) On general minimax theorems. *Pacific J. Math.* **8**, 171–176.
- [29] H. SPOHN (2014) KPZ scaling theory and the semidiscrete directed polymer model. In: *Random Matrix Theory, Interacting Particle Systems, and Integrable Systems*. Math. Sci. Res. Inst. Publ. **65**, Cambridge Univ. Press, New York, 483–493.