The Annals of Probability
2013, Vol. 41, No. 6, 4248–4286
DOI: 10.1214/12-AOP768

LARGE DEVIATION RATE FUNCTIONS FOR THE PARTITION FUNCTION IN A LOG-GAMMA DISTRIBUTED RANDOM POTENTIAL

BY NICOS GEORGIOU AND TIMO SEPPÄLÄINEN[1]

University of Utah and University of Wisconsin–Madison

We study right tail large deviations of the logarithm of the partition function for directed lattice paths in i.i.d. random potentials. The main purpose is the derivation of explicit formulas for the $1+1$-dimensional exactly solvable case with log-gamma distributed random weights. Along the way we establish some regularity results for this rate function for general distributions in arbitrary dimensions.

1. Introduction. We study a version of the model called *directed polymer in a random environment* where a fluctuating path is coupled with a random environment. This model was introduced in the statistical physics literature in [16] and early mathematically rigorous work followed in [3, 17]. We consider directed paths in the nonnegative orthant $\mathbb{Z}_+^d$ of the d-dimensional integer lattice. The paths are allowed nearest-neighbor steps oriented along the coordinate axes. A random weight $\omega(\mathbf{u})$ is attached to each lattice point $\mathbf{u} \in \mathbb{Z}_+^d$. Together the weights form the *environment* $\omega = \{\omega(\mathbf{u}) : \mathbf{u} \in \mathbb{Z}_+^d\}$. The space of environments is denoted by Ω. $\mathbb{P}$ is a probability measure on Ω under which the weights $\{\omega(\mathbf{u})\}$ are i.i.d. random variables.

For $\mathbf{v}, \mathbf{u} \in \mathbb{Z}_+^d$ such that $\mathbf{v} \le \mathbf{u}$ (coordinatewise ordering), the set of admissible paths from $\mathbf{v}$ to $\mathbf{u}$ with $|\mathbf{u} - \mathbf{v}|_1 = m$ is

$$\begin{aligned} \Pi_{\mathbf{v},\mathbf{u}} = \big\{x_\centerdot = \{\mathbf{v} = x_0, x_1, \dots, x_m = \mathbf{u}\} : \forall k, x_k \in \mathbb{Z}_+^d \text{ and} \\ x_{k+1} - x_k \in \{\mathbf{e}_i : 1 \le i \le d\}\big\}, \end{aligned} \tag{1.1}$$

where e_i is the ith standard basis vector of $\mathbb{R}^d$. The *point-to-point partition function* is

$$Z_{\mathbf{v},\mathbf{u}} = \sum_{x_\centerdot \in \Pi_{\mathbf{v},\mathbf{u}}} e^{\sum_{j=1}^m \omega(x_j)}. \tag{1.2}$$

Received November 2011; revised April 2012.
[1]Supported in part by NSF Grant DMS-10-03651 and by the Wisconsin Alumni Research Foundation.
MSC2010 subject classifications. Primary 60K37; secondary 60K35, 60F10.
Key words and phrases. Directed polymer in random environment, partition function, large deviations, random walk in random potential.

This is the normalization factor in the *quenched polymer distribution*

$$Q_{\mathbf{v},\mathbf{u}}(x_{\bullet}) = Z_{\mathbf{v},\mathbf{u}}^{-1} \prod_{j=1}^{m} e^{\omega(x_j)}, \tag{1.3}$$

which is a probability distribution on the paths in the set $\Pi_{\mathbf{v},\mathbf{u}}$. When paths start at the origin ($\mathbf{v} = \mathbf{0}$), we drop $\mathbf{v}$ from the notation; $Z_{\mathbf{u}} = Z_{\mathbf{0},\mathbf{u}}$ and $\Pi_{\mathbf{u}} = \Pi_{\mathbf{0},\mathbf{u}}$. Note that the weight at the starting point x_0 was not included in the sum in the exponent in (1.2). This makes no difference for the results. Sometimes it is convenient to include this weight, and then we write $Z_{\mathbf{v},\mathbf{u}}^{\square} = e^{\omega(\mathbf{v})} Z_{\mathbf{v},\mathbf{u}}$ where the superscript $\square$ reminds us that all weights in the rectangle are included.

In the polymer model one typically studies fluctuations of the path and fluctuations of $\log Z_{\mathbf{u}}$. This paper considers only $\log Z_{\mathbf{u}}$. Specifically our main object of interest is the right tail large deviation rate function

$$J_{\mathbf{u}}(r) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \ge nr\} \tag{1.4}$$

for $\mathbf{u} \in \mathbb{R}_{+}^{d}$, $r \in \mathbb{R}$. Throughout we denote the floor of a vector as $\lfloor n\mathbf{y}\rfloor = (\lfloor ny_1\rfloor, \lfloor ny_2\rfloor, \dots, \lfloor ny_d\rfloor)$. This function J exists very generally for superadditivity reasons, and in Section 3 we establish some of its regularity properties.

The focus of the paper is an exactly solvable case where $d = 2$ and $-\omega(\mathbf{u})$ is log-gamma distributed. By "exactly solvable" we mean that special properties of the log-gamma case permit explicit computations, such as a formula for the limiting point-to-point free energy

$$p(\mathbf{y}) = \lim_{n\to\infty} n^{-1} \log Z_{\lfloor n\mathbf{y}\rfloor}, \qquad \mathbb{P}\text{-a.s.} \tag{1.5}$$

and fluctuation exponents [31]. In the same spirit, in this paper we compute explicit formulas for the rate function J and other related quantities in the context of the $1+1$-dimensional log-gamma polymer.

One can also consider point-to-line partition functions over all directed paths of a fixed length. For $m \in \mathbb{N}$ the partition function is defined by

$$Z_m^{\text{line}} = \sum_{\mathbf{u}\in\mathbb{Z}_{+}^{d} : |\mathbf{u}|_1 = m} Z_{\mathbf{u}}. \tag{1.6}$$

Due to the $n^{-1}\log$ in front, in the results we look at Z_m^{line} behaves like the maximal $Z_{\mathbf{u}}$ over $|\mathbf{u}|_1 = m$.

Some comments are in order.

There are currently three known exactly solvable directed polymer models, all in $1+1$ dimensions: the two with a discrete aspect are (i) the log-gamma model introduced in [31], and (ii) a model introduced in [27] where the random environment is a collection of Brownian motions. Some fluctuation exponents were derived for the second model in [32], and it has been further studied in [26] via a connection with the quantum Toda lattice. This Brownian model possesses structures similar to those in the log-gamma model, so we expect that the results of the present paper could be reproduced for the Brownian model.

The third exactly solvable model is the continuum directed random polymer [1] that is expected to be a universal scaling limit for a large class of polymer models; see [10] for a recent review.

Usually the directed lattice polymer model is placed in a space–time picture where the paths are oriented in the time direction. (See articles and lectures [5, 6, 8, 13] for recent results and reviews of the general case.) In two dimensions (1 time + 1 space dimension), the space–time picture is the same as our purely spatial picture, up to a 45° rotation of the lattice and a change of lattice indices. The temporal aspect is not really present in our work. So we have not separated a time dimension, but simply regard the paths as directed lattice paths.

Another standard feature of directed polymers that we have omitted is the inverse temperature parameter $\beta \in (0, \infty)$ that appears as a multiplicative constant in front of the weights: $Z^{\beta}_{\mathbf{v},\mathbf{u}} = \sum_{x_\bullet \in \Pi_{\mathbf{v},\mathbf{u}}} \exp\{\beta \sum_{j=1}^{m} \omega(x_j)\}$. For a fixed weight distribution, β modulates the strength of the coupling between the walk and the environment. It is known that in dimension $1 + 3$ and higher, there can be a phase transition. By contrast, in low dimensions ($1 + 1$ and $1 + 2$), the model is in the so-called strong coupling regime for all $0 < \beta < \infty$ [7, 21]. The β parameter plays no role in the present work and has a fixed value $\beta = 1$. This is the unique β value that turns the log-gamma model into an exactly solvable model.

The techniques of the current paper are entirely probabilistic and rely on the stationary version of the log-gamma model. It can be expected that as a combinatorial approach to this model, fully developed [11], more complete results and alternative proofs for the present results can be found.

Earlier literature. Precise large deviation rate functions for $\log Z$ in the case of directed polymers have not been derived in the past. The strongest concentration inequalities can be found in recent references [9, 22, 33]. The normalization of the left tail varies with the distribution of the weights as demonstrated by [2], but the right tails have the same normalization n. Carmona and Hu [4] have some bounds on the left tail of $\log Z$ in Gaussian environments in dimensions $1 + 3$ and higher and for small enough β. Similar bounds were proved later in [24] for bounded environments using concentration inequalities for product measures.

For the exactly solvable zero-temperature models (i.e., last passage percolation models), large deviation principles have been proved. For the longest increasing path among planar Poisson points, an LDP for the length resulted from a combination of articles [14, 20, 23, 30]. These results came before the advent of determinantal techniques. For the corner growth model with geometric and exponential weights [18] derived an LDP in addition to the Tracy–Widom limit. An earlier right tail LDP appeared in [29].

Notation. We collect some notation and conventions here for easy reference. $\mathbb{N}$ is for positive integers, $\mathbb{Z}_+$ for nonnegative integer, $\mathbb{R}_+$ for nonnegative real numbers and $\mathbb{R}_+^d$ is the set of all vectors with nonnegative real coordinates. Vector notation: elements of $\mathbb{R}^d$ and $\mathbb{Z}^d$ are $\mathbf{v} = (v_1, v_2, \dots, v_d)$. Coordinatewise ordering $\mathbf{v} \le \mathbf{u}$ means $v_1 \le u_1, v_2 \le u_2, \dots, v_d \le u_d$. Particular vectors are $\mathbf{1} = (1, 1, \dots, 1)$ and $\mathbf{0} = (0, 0, \dots, 0)$. $\lfloor \mathbf{y} \rfloor = (\lfloor y_1 \rfloor, \lfloor y_2 \rfloor, \dots, \lfloor y_d \rfloor)$ where $\lfloor y \rfloor = \max\{n \in \mathbb{Z} : n \le y\}$ is the integer part of $y \in \mathbb{R}$. The ℓ^1 norm on $\mathbb{R}^d$ is $|\mathbf{v}|_1 = |v_1| + \cdots + |v_d|$.

The convex dual of a function $f : \mathbb{R} \to (-\infty, \infty]$ is $f^*(y) = \sup_{x \in \mathbb{R}}\{xy - f(x)\}$, and $f = f^{**}$ if and only if f is convex and lower semicontinuous. We refer to [28] for basic convex analysis.

The partition function Z does not include the weight of the initial point of the paths, while $Z^\square$ does. In two dimensions we write $Z_{m,n} = Z_{(m,n)}$.

The usual gamma function is $\Gamma(\mu) = \int_0^\infty x^{\mu-1} e^{-x}\,dx$ for $\mu > 0$. The digamma and trigamma functions are $\Psi_0 = \Gamma'/\Gamma$ and $\Psi_1 = \Psi_0'$. On $(0, \infty)$ Ψ_0 is increasing and concave and Ψ_1 decreasing, positive and convex, with $-\Psi_0(0+) = \Psi_1(0+) = \infty$.

2. Large deviations for the log-gamma model.

2.1. *The log-gamma model with i.i.d. weights.* In this section we specialize to $d = 2$ dimensions and the log-gamma distributed weights. Fix a positive real parameter μ. This parameter remains fixed through this entire section, and hence is omitted from most notation. In the log-gamma case we prefer to switch to multiplicative variables. So the weight at point $(i, j) \in \mathbb{Z}_+^2$ is $Y_{i,j} = e^{\omega(i,j)}$ where the reciprocal Y^{-1} has Gamma(μ) distribution. Explicitly,

$$\mathbb{P}\{Y^{-1} \ge s\} = \Gamma(\mu)^{-1} \int_s^\infty x^{\mu-1} e^{-x}\,dx \qquad \text{for } s \in \mathbb{R}_+. \tag{2.1}$$

As above, we write Y for a generic random variable distributed as $Y_{i,j}$. The digamma and trigamma functions give the mean and variance, $\mathbb{E}(\log Y) = -\Psi_0(\mu)$ and $\mathbb{V}\mathrm{ar}(\log Y) = \Psi_1(\mu)$.

The logarithmic moment generating function (l.m.g.f.) of $\omega = \log Y$ is

$$M_\mu(\xi) = \log \mathbb{E}\big(e^{\xi \log Y}\big) = \begin{cases} \log\Gamma(\mu - \xi) - \log\Gamma(\mu), & \xi \in (-\infty, \mu), \\ \infty, & \xi \in [\mu, \infty). \end{cases} \tag{2.2}$$

The point-to-point partition function for directed paths from $(0, 0)$ to (m, n) is

$$Z_{m,n} = \sum_{x_\cdot \in \Pi_{(m,n)}} \prod_{j=1}^{m+n} Y_{x_j}. \tag{2.3}$$

Note that we simplified notation by dropping the parentheses: $Z_{m,n} = Z_{(m,n)}$. For $(s,t) \in \mathbb{R}_+^2$ the limiting free energy density exists by superadditivity,

$$p(s,t) = \lim_{n\to\infty} n^{-1} \log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor}, \qquad \mathbb{P}\text{-a.s.} \tag{2.4}$$

The limit is a finite constant. We begin by giving its exact value.

THEOREM 2.1. *For $(s,t) \in \mathbb{R}_+^2$ and $\mu \in (0,\infty)$, the limiting free energy density (2.4) is given by*

$$p(s,t) = \inf_{0<\rho<\mu} \{-s\Psi_0(\rho) - t\Psi_0(\mu-\rho)\}. \tag{2.5}$$

The value $p(s,t)$ was already derived in [31] but the proof was buried among estimates for fluctuation exponents. In Section 4 we sketch an elementary approach that utilizes special features of the log-gamma model. For the other explicitly solvable $1+1$-dimensional polymer with Brownian environment, Moriarty and O'Connell [25] computed the limiting free energy with a very different large deviation approach.

The next result is a large deviation principle (LDP) for $\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ under normalization n. The rate function is

$$I_{s,t}(r) = \begin{cases} \sup_{\xi\in[0,\mu)} \Big\{ r\xi - \inf_{\theta\in(\xi,\mu)} \big(tM_\theta(\xi) - sM_{\mu-\theta}(-\xi)\big)\Big\}, & r \ge p(s,t), \\ \infty, & r < p(s,t). \end{cases} \tag{2.6}$$

On the boundary ($s=0$ or $t=0$), the result reduces to i.i.d. large deviations, so we only consider (s,t) in the interior of the quadrant.

THEOREM 2.2. *Let $Y^{-1} \sim \mathrm{Gamma}(\mu)$ as in (2.1) and $(s,t) \in (0,\infty)^2$. Then the distributions of $n^{-1}\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ satisfy a LDP with normalization n and rate function $I_{s,t}$. Explicitly, these bounds hold for any open set G and any closed set F in $\mathbb{R}$:*

$$\varlimsup_{n\to\infty} n^{-1} \log \mathbb{P}\{n^{-1}\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor} \in F\} \le -\inf_{r\in F} I_{s,t}(r) \tag{2.7}$$

and

$$\varliminf_{n\to\infty} n^{-1} \log \mathbb{P}\{n^{-1}\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor} \in G\} \ge -\inf_{r\in G} I_{s,t}(r). \tag{2.8}$$

On $[p(s,t),\infty)$ the rate function $I_{s,t}$ is finite, strictly increasing, continuous and convex. In particular, the unique zero of $I_{s,t}(r)$ is at $r = p(s,t)$. The right tail rate defined in (1.4) is given by

$$J_{s,t}(r) = \begin{cases} 0, & r \in (-\infty, p(s,t)], \\ I_{s,t}(r), & r \in [p(s,t),\infty). \end{cases} \tag{2.9}$$

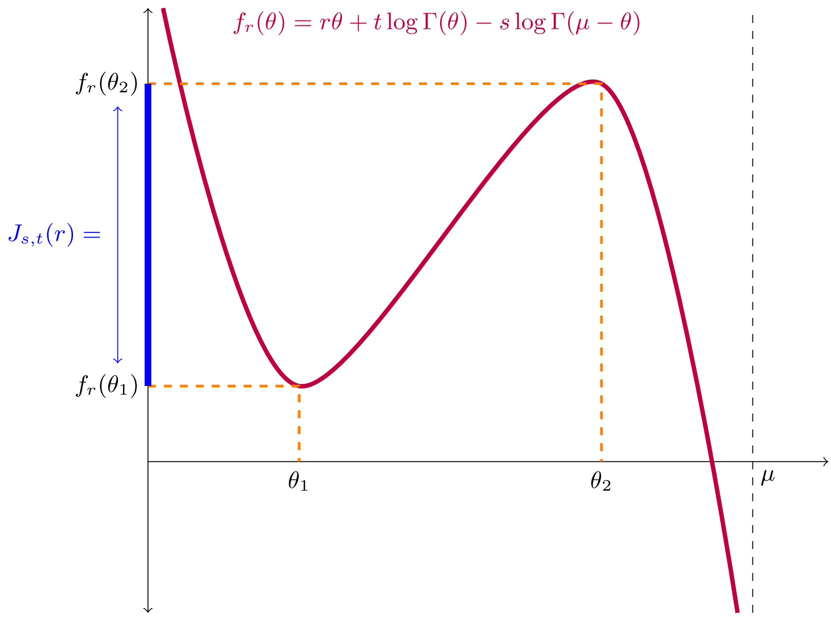

FIG. 1. *Graphical representation of the solution to the variational problem* (2.6) *that gives the rate function* $J_{s,t}(r) = f_r(\theta_2) - f_r(\theta_1)$. *The curve* $f_r(\theta)$ *has the same general shape as long as* $r > p(s,t)$.

REMARK 2.3. From a computational point of view, the solution to the variational problem in (2.6) can be computed by

$$I_{s,t}(r) = \sup_{0<\theta<\mu}\Big\{f_r(\theta) - \inf_{0<z\le\theta} f_r(z)\Big\} = f_r(\theta_2) - f_r(\theta_1),$$

where

$$f_r(\theta) = r\theta + t\log\Gamma(\theta) - s\log\Gamma(\mu-\theta),$$

and for any $r > p(s,t)$, $0 < \theta_1 < \theta_2 < \mu$ are the solutions to the equation $\frac{d}{d\theta} f_r(\theta) = 0$. (See Figure 1.) This again implies that the rate function is strictly positive as long as $r > p(s,t)$.

REMARK 2.4. We do not address the precise large deviations in the left tail, that is, in the range $r < p(s,t)$. We expect the correct normalization to be n^2. (Personal communication from I. Ben-Ari.) Presently we do not have a technique for computing the rate function in that regime. We include the trivial part $I_{s,t}(r) = \infty$ for $r < p(s,t)$ in the theorem so that we can compute the limiting l.m.g.f. by a straightforward application of Varadhan's theorem.

Define for $\xi \in \mathbb{R}$,

$$\Lambda_{s,t}(\xi) = \lim_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor}}. \tag{2.10}$$

COROLLARY 2.5. *Let $\xi \in \mathbb{R}$. Then the limit in* (2.10) *exists and is given by*

$$\Lambda_{s,t}(\xi) = I^*_{s,t}(\xi) = \begin{cases} p(s,t)\xi, & \xi < 0, \\ \inf_{\theta \in (\xi,\mu)} \{tM_\theta(\xi) - sM_{\mu-\theta}(-\xi)\}, & 0 \le \xi < \mu, \\ \infty, & \xi \ge \mu. \end{cases} \tag{2.11}$$

REMARK 2.6. Symmetry of $\Lambda_{s,t}$ in (s,t) is clear from (2.10) but not immediately obvious in the $0 \le \xi < \mu$ case of (2.11). It turns out that if $s \le t$ the infimum is achieved at a unique $\theta_0 \in [(\mu + \xi)/2, \mu)$, and then for $\Lambda_{t,s}(\xi)$, the same infimum is uniquely achieved at $\theta_1 = \mu + \xi - \theta_0 \in (\xi, (\mu + \xi)/2]$. In the case $s = t$ a simple formula arises: $\Lambda_{t,t}(\xi) = 2t(\log\Gamma(\frac{\mu-\xi}{2}) - \log\Gamma(\frac{\mu+\xi}{2}))$.

REMARK 2.7. The first case of (2.6) gives $I_{s,t}$ as the dual $\Lambda^*_{s,t}$, and the reader may wonder whether this is the logic of the proof of the LDP. It is not, for we have no direct way to compute $\Lambda_{s,t}$. Instead, Theorem 2.2 is first proved in an indirect manner via the stationary model described in the next subsection, and then $\Lambda_{s,t}$ is derived by Varadhan's theorem.

Let us also record the result for the point-to-line case. It behaves like the point-to-point case along the diagonal.

COROLLARY 2.8. *Let $Y^{-1} \sim \mathrm{Gamma}(\mu)$ as in* (2.1) *and $s > 0$. Then the distributions of $\log Z^{\mathrm{line}}_{\lfloor ns \rfloor}$ satisfy an LDP with normalization n and rate function $I_{s/2,s/2}$.*

REMARK 2.9. For $\varepsilon > 0$ and $r = p(s,t) + \varepsilon$, one can show after some calculus that there exists a nonzero constant $C = C_{s,t}(\mu)$ so that

$$I_{s,t}(r) = C\varepsilon^{3/2} + o(\varepsilon^{3/2}).$$

This suggests that $\mathbb{V}\mathrm{ar}(\log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor})$ is of order $n^{2/3}$. Rigorous upper bounds on the moments $\mathbb{E}|\log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor} - np(s,t)|^p$ for $1 \le p < 3/2$ can be found in [31], Theorem 2.4.

We computed the precise value of the constant C for the point-to-line rate function,

$$I_{1,1}(r) = \frac{4}{3} \frac{1}{\sqrt{|\Psi_2(\mu/2)|}} \varepsilon^{3/2} + o(\varepsilon^{3/2}), \tag{2.12}$$

where $\Psi_2 = \Psi_0''$.

2.2. *The stationary log-gamma model.* Next we consider the log-gamma model in a stationary situation that is special to this choice of distribution. Working with the stationary case is the key to explicit computations, including all the previous results, and provides some explanation for the formulas that arose for $I_{s,t}$ and $\Lambda_{s,t}$ in (2.6) and (2.11).

The stationary model is created by appropriately altering the distributions of the weights on the boundaries of the quadrant $\mathbb{Z}_+^2$. We continue to use the parameter $\mu \in (0, \infty)$ fixed at the beginning of this section, and we introduce a second parameter $\theta \in (0, \mu)$. Let the collection of independent weights $\{U_{i,0}, V_{0,j}, Y_{i,j} : i, j \in \mathbb{N}\}$ have the following marginal distributions:

$$U_{i,0}^{-1} \sim \mathrm{Gamma}(\theta), \qquad V_{0,j}^{-1} \sim \mathrm{Gamma}(\mu - \theta) \quad \text{and} \tag{2.13}$$
$$Y_{i,j}^{-1} \sim \mathrm{Gamma}(\mu).$$

Define the partition function $Z_{m,n}^{(\theta)}$ by (2.3) with the following weights: at the origin $Y_{0,0} = 1$, on the x-axis $Y_{i,0} = U_{i,0}$, on the y-axis $Y_{0,j} = V_{0,j}$, and in the bulk the weights $\{Y_{i,j} : i, j \in \mathbb{N}\}$ are i.i.d. $\mathrm{Gamma}(\mu)^{-1}$ as before. Equivalently, we can decompose the stationary partition function $Z_{m,n}^{(\theta)}$ according to the exit point of the path from the boundary

$$Z_{m,n}^{(\theta)} = \sum_{k=1}^{m} \left(\prod_{i=1}^{k} U_{i,0} \right) Z_{(k,1),(m,n)}^{\square} + \sum_{\ell=1}^{n} \left(\prod_{j=1}^{\ell} V_{0,j} \right) Z_{(1,\ell),(m,n)}^{\square}. \tag{2.14}$$

The symbols $U_{i,0}$ and $V_{0,j}$ were at first introduced for the boundary weights to highlight the change of distribution. Next let us define for all $(i, j) \in \mathbb{Z}_+^2 \setminus \{(0, 0)\}$,

$$U_{i,j} = \frac{Z_{i,j}^{(\theta)}}{Z_{i-1,j}^{(\theta)}} \quad \text{and} \quad V_{i,j} = \frac{Z_{i,j}^{(\theta)}}{Z_{i,j-1}^{(\theta)}}. \tag{2.15}$$

Note that this property was already built into the boundaries because, for example, $Z_{i,0}^{(\theta)} = U_{1,0} \cdots U_{i,0}$. The key result that allows explicit calculations for this model is the following.

PROPOSITION 2.10. *For each $(i, j) \in \mathbb{Z}_+^2 \setminus \{(0, 0)\}$, we have the following marginal distributions*: $U_{i,j}^{-1} \sim \mathrm{Gamma}(\theta)$ *and* $V_{i,j}^{-1} \sim \mathrm{Gamma}(\mu - \theta)$. *For any fixed $n \in \mathbb{Z}_+$, the variables $\{U_{i,n} : i \in \mathbb{N}\}$ are i.i.d., and for any fixed $m \in \mathbb{Z}_+$, the variables $\{V_{m,j} : j \in \mathbb{N}\}$ are i.i.d.*

This is a special case of Theorem 3.3 in [31], where the independence of these weights along more general down-right lattice paths is established. Proposition 2.10 is the only result from [31] that we use. It follows in an elementary fashion from the properties of the gamma distribution.

As an immediate application we can write

$$n^{-1}\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} = n^{-1}\sum_{j=1}^{\lfloor nt\rfloor}\log V_{0,j} + n^{-1}\sum_{i=1}^{\lfloor ns\rfloor}\log U_{i,\lfloor nt\rfloor} \tag{2.16}$$

as a sum of two sums of i.i.d. variables, and from this compute

$$\mathbb{E}\big(\log Z^{(\theta)}_{m,n}\big) = m\mathbb{E}(\log U) + n\mathbb{E}(\log V) = -m\Psi_0(\theta) - n\Psi_0(\mu-\theta) \tag{2.17}$$

and obtain the law of large numbers,

$$n^{-1}\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \to p^{(\theta)}(s,t) = -s\Psi_0(\theta) - t\Psi_0(\mu-\theta), \qquad \mathbb{P}\text{-a.s.} \tag{2.18}$$

Note that the two sums on the right-hand side of (2.16) are not independent of each other. In fact, they are so strongly negatively correlated that the variance of their sum is of order $n^{2/3}$ [31]. Comparison of (2.5) and (2.18) reveals a variational principle at work: $p(s,t)$ is the minimal free energy of a stationary system with bulk parameter μ.

Instead of the right tail large deviation rate function, we give the asymptotic l.m.g.f. in the next result. Define

$$\Lambda_{\theta,(s,t)}(\xi) = \lim_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}. \tag{2.19}$$

THEOREM 2.11. *Let $s,t\ge 0$ and $0<\theta<\mu$. Then the limit in* (2.19) *exists for $\xi\ge 0$ and is given by*

$$\Lambda_{\theta,(s,t)}(\xi) = \begin{cases} \max\{sM_\theta(\xi) - tM_{\mu-\theta}(-\xi),\, tM_{\mu-\theta}(\xi) - sM_\theta(-\xi)\}, \\ \qquad\qquad 0\le\xi<\theta\wedge(\mu-\theta) \\ \infty, \qquad \xi\ge\theta\wedge(\mu-\theta). \end{cases} \tag{2.20}$$

REMARK 2.12. Let the parameters $0<\theta<\mu$ be given. The *characteristic direction* is the choice

$$(s,t) = c\big(\Psi_1(\mu-\theta), \Psi_1(\theta)\big) \qquad \text{for a constant } c>0. \tag{2.21}$$

With this choice the variance of $\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ is of order $n^{2/3}$, while in other directions the fluctuations of $\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ have order of magnitude $n^{1/2}$ and they are asymptotically Gaussian [31]. By this token, we would expect the large deviations in the characteristic situation to be unusual, while in the off-characteristic directions we would expect the more typical large deviations of order e^{-n} in both tails. In Lemma 4.2(b) we give a bound on the left tail that indicates superexponential decay under (2.21). This also implies that if (2.21) holds, then formula (2.20) can be complemented with the case $\Lambda_{\theta,(s,t)}(\xi) = p^{(\theta)}(s,t)\xi$ for $\xi\le 0$. Presently we do not have further information about these large deviations.

REMARK 2.13. If the two sums in (2.16) were independent we would have $\Lambda_{\theta,(s,t)}(\xi) = sM_\theta(\xi) + tM_{\mu-\theta}(\xi)$. Obviously (2.20) reflects the strong negative correlation of these sums, but currently we do not have a good explanation (besides the proof!) for the formula that arises.

The maximum in (2.20) comes from the choice of the first step of the path: either horizontal or vertical. Corresponding to this choice, define partition functions

$$Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} = \sum_{k=1}^{\lfloor ns\rfloor}\left(\prod_{i=1}^{k} U_{i,0}\right) Z^{\square}_{(k,1),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \tag{2.22}$$

and

$$Z^{(\theta),\mathrm{ver}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} = \sum_{\ell=1}^{\lfloor nt\rfloor}\left(\prod_{j=1}^{\ell} V_{0,j}\right) Z^{\square}_{(1,\ell),(\lfloor ns\rfloor,\lfloor nt\rfloor)}, \tag{2.23}$$

together with l.m.g.f.'s

$$\Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) = \lim_{n\to\infty} n^{-1}\log \mathbb{E}e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}} \tag{2.24}$$

and

$$\Lambda^{\mathrm{ver}}_{\theta,(s,t)}(\xi) = \lim_{n\to\infty} n^{-1}\log \mathbb{E}e^{\xi\log Z^{(\theta),\mathrm{ver}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}. \tag{2.25}$$

Then $Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} = Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} + Z^{(\theta),\mathrm{ver}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ leads to

$$\Lambda_{\theta,(s,t)}(\xi) = \Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) \vee \Lambda^{\mathrm{ver}}_{\theta,(s,t)}(\xi), \tag{2.26}$$

which is the starting point for the proof of (2.20).

The horizontal and vertical partition functions are in some sense between the stationary one and the one from (2.3) with i.i.d. weights. It turns out that these intermediate partition functions behave either like the stationary one or like the i.i.d. one, with a sharp transition in between, and this holds both at the level of the limiting free energy density and the l.m.g.f. Let us focus on the horizontal case, the vertical case being the same after the swap $s \leftrightarrow t$ and $\theta \leftrightarrow \mu - \theta$.

Qualitatively, with t fixed, when s is large $Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ behaves like $Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$, and when s is small $Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ behaves like $Z_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ from (2.3). The conditions for the transitions are the following:

$$s\Psi_1(\theta) \ge t\Psi_1(\mu-\theta) \tag{2.27}$$

and

$$s\bigl(\Psi_0(\theta) - \Psi_0(\theta-\xi)\bigr) \ge t\bigl(\Psi_0(\mu-\theta+\xi) - \Psi_0(\mu-\theta)\bigr). \tag{2.28}$$

By the concavity of Ψ_0 and the fact that $\Psi_1 = \Psi_0'$, (2.27) implies (2.28) for all $\xi \geq 0$. Assuming the limit exists for the moment, define

$$p^{(\theta),\mathrm{hor}}(s,t) = \lim_{n\to\infty} n^{-1} \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}. \tag{2.29}$$

In this next theorem the functions $p(s,t)$ and $\Lambda_{s,t}(\xi)$ are the ones defined by (2.5) and (2.11).

THEOREM 2.14. *Let $s,t \geq 0$, $0 < \theta < \mu$ and $0 \leq \xi < \theta$.*

(a) *The limit in* (2.29) *exists and is given by*

$$p^{(\theta),\mathrm{hor}}(s,t) = \begin{cases} p^{(\theta)}(s,t), & \text{if (2.27) holds,} \\ p(s,t), & \text{if (2.27) fails.} \end{cases} \tag{2.30}$$

(b) *The limit in* (2.24) *exists and is given by*

$$\Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) = \begin{cases} sM_\theta(\xi) - tM_{\mu-\theta}(-\xi), & \text{if (2.28) holds,} \\ \Lambda_{s,t}(\xi), & \text{if (2.28) fails.} \end{cases} \tag{2.31}$$

REMARK 2.15. We saw in (2.5) that the limiting free energy $p(s,t)$ of the i.i.d. model is the minimal free energy of the stationary models with the same bulk parameter μ. This link does *not* extend to the l.m.g.f.'s: for $0 < \xi < \mu$, $\Lambda_{s,t}(\xi) < \Lambda_{\theta,(s,t)}(\xi)$ for all $\theta \in (0,\mu)$. We observe this at the end of the proof of Theorem 2.11 in Section 5.

3. The right tail rate function in the general case. The proofs of the results for the log-gamma model utilize regularity properties of the rate function J of (1.4). These properties can be proved in some degree of generality, and we do so in this section. So now we consider

$$Z_{\mathbf{u}} = \sum_{x_\bullet \in \Pi_{\mathbf{u}}} e^{\sum_{j=1}^{|u|_1} \omega(x_j)} \tag{3.1}$$

as defined in the Introduction, with $\mathbf{u} \in \mathbb{Z}_+^d$, general $d \geq 2$, and general i.i.d. weights $\{\omega(\mathbf{u})\}$.

We assume

$$\exists \xi > 0 \text{ such that } \mathbb{E}\big(e^{\xi|\omega(\mathbf{u})|}\big) < \infty. \tag{3.2}$$

This guarantees the existence of a Cramér large deviation rate function defined by

$$I(r) = -\lim_{\varepsilon\to 0}\lim_{n\to\infty} n^{-1} \log \mathbb{P}\left\{ n^{-1} \sum_{i=1}^{n} \omega(\mathbf{u}_i) \in (r-\varepsilon, r+\varepsilon) \right\}. \tag{3.3}$$

(Above $\{\mathbf{u}_j\}$ are any distinct lattice points.) We state first the existence theorem for the limiting point-to-point free energy density. We omit the proof because similar superadditive and approximation arguments appear elsewhere in our paper, and refer to [15]. Let us also point out that assumption (3.2) is unnecessarily strong for this existence result, but our objective here is not to optimize on this point.

THEOREM 3.1. *Assume* (3.2). *There exists an event* $\Omega_0 \subseteq \Omega$ *of full* $\mathbb{P}$-*probability on which the convergence*

$$p(\mathbf{y}) = \lim_{n\to\infty} n^{-1} \log Z_{\lfloor n\mathbf{y}\rfloor} \tag{3.4}$$

happens simultaneously for all $\mathbf{y} \in \mathbb{R}_+^d$. *Limit* (3.4) *holds also in* $L^1(\mathbb{P})$. *As a function of* $\mathbf{y}$, p *is concave and continuous on* $\mathbb{R}_+^d$.

Next the right-tail LDP. To avoid issues of vanishing probabilities and infinite values of the rate, we make the following further assumption:

$$\forall r < \infty \qquad \mathbb{P}\{\omega(\mathbf{0}) > r\} > 0. \tag{3.5}$$

THEOREM 3.2. *Assume* (3.2) *and* (3.5). *Then for* $\mathbf{u} \in \mathbb{R}_+^d \setminus \{\mathbf{0}\}$ *and* $r \in \mathbb{R}$, *the following* $\mathbb{R}_+$-*valued limit exists*:

$$J_{\mathbf{u}}(r) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \geq nr\}. \tag{3.6}$$

As a function of $(\mathbf{u}, r)$, J *is convex and continuous on* $(\mathbb{R}_+^d \setminus \{\mathbf{0}\}) \times \mathbb{R}$. $J_{\mathbf{u}}(r) = 0$ *if and only if* $r \leq p(\mathbf{u})$.

Let us also remark that the weight $\omega(\mathbf{0})$ at the origin is immaterial: the limit is the same for $Z^{\square}$, so for $\mathbf{u} \in \mathbb{R}_+^d \setminus \{\mathbf{0}\}$ and $r \in \mathbb{R}$,

$$J_{\mathbf{u}}(r) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z^{\square}_{\lfloor n\mathbf{u}\rfloor} \geq nr\}. \tag{3.7}$$

We observe this at the end of the proof of Theorem 3.2.

With a further assumption on the Cramér rate function of the weight distribution defined in (3.3), we can extend the continuity of $J_{\mathbf{u}}$ to $\mathbf{u} = 0$:

$$\alpha_\infty = \lim_{x\nearrow\infty} x^{-1} I(x) < \infty. \tag{3.8}$$

Equation (3.5) is equivalent to requiring that $I(x) < \infty$ for all large enough x, so of course (3.8) requires (3.5). The constant α_∞ is the limiting slope of I at ∞ which exists by convexity. When assumption (3.8) is in force we define

$$J_{\mathbf{0}}(r) = \begin{cases} 0, & r \leq 0, \\ \alpha_\infty r, & r \geq 0. \end{cases} \tag{3.9}$$

THEOREM 3.3. *Under assumptions* (3.2) *and* (3.8), *and with* $J_{\mathbf{0}}$ *defined by* (3.9), $J_{\mathbf{u}}(r)$ *is finite and continuous on* $\mathbb{R}_+^d \times \mathbb{R}$.

REMARK 3.4. Assumption (3.8) is in particular valid for the log-gamma model. For $Y^{-1} \sim \text{Gamma}(\mu)$ the Cramér rate function for $\omega = \log Y$ is

$$I_\mu(r) = -r\Psi_0^{-1}(-r) - \log\Gamma\big(\Psi_0^{-1}(-r)\big) + \mu r + \log\Gamma(\mu), \qquad r \in \mathbb{R}. \tag{3.10}$$

The limiting slope on the right is $\alpha_\infty = \mu$, while the limiting slope on the left would be $\lim_{r\to-\infty} I'(r) = -\infty$. In this case $J_{\mathbf{0}}(r)$ is also the "rate function" for the single weight at the origin

$$J_{\mathbf{0}}(r) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Y \ge nr\}. \tag{3.11}$$

The remainder of this section proves Theorems 3.2 and 3.3, and then we prove two further lemmas for later use.

PROOF OF THEOREM 3.2. For $m, n \in \mathbb{R}_+$, let $\mathbf{x}_{m,n} \in \{0,1\}^d$ so that $\lfloor (m+n)\mathbf{u}\rfloor = \lfloor m\mathbf{u}\rfloor + \lfloor n\mathbf{u}\rfloor + \mathbf{x}_{m,n}$. By superadditivity, independence and shift invariance,

$$\begin{aligned} &\mathbb{P}\bigl\{\log Z_{\lfloor (m+n)\mathbf{u}\rfloor} \ge (m+n)r\bigr\} \\ &\qquad \ge \mathbb{P}\{\log Z_{\lfloor m\mathbf{u}\rfloor} \ge mr\}\mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \ge nr\}\mathbb{P}\{\log Z_{\mathbf{x}_{m,n}} \ge 0\}. \end{aligned} \tag{3.12}$$

By assumption (3.5) there is a uniform lower bound $\mathbb{P}\{\log Z_{\mathbf{x}_{m,n}} \ge 0\} \ge \rho > 0$. Thus $t(n) = \log \mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \ge nr\}$ is superadditive with a small uniformly bounded correction. Assumption (3.5) implies that $t(n) > -\infty$ for all $n \ge n_0$. Consequently by superadditivity the rate function

$$J_{\mathbf{u}}(r) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \ge nr\} \tag{3.13}$$

exists for $\mathbf{u} = (u_1, \dots, u_d) \in \mathbb{R}_+^d$ and $r \in \mathbb{R}$. The limit in (3.13) holds also as $n \to \infty$ through real values, not just integers.

Similarly we get convexity of J in $(\mathbf{u}, r)$. Let $\lambda \in (0,1)$ and assume $(\mathbf{u}, r) = \lambda(\mathbf{u}_1, r_1) + (1-\lambda)(\mathbf{u}_2, r_2)$. Then

$$\begin{aligned} &n^{-1} \log \mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \ge nr\} \\ &\quad \ge \lambda(\lambda n)^{-1} \log \mathbb{P}\{\log Z_{\lfloor n\lambda\mathbf{u}_1\rfloor} \ge n\lambda r_1\} \\ &\qquad + (1-\lambda)\bigl((1-\lambda)n\bigr)^{-1} \log \mathbb{P}\bigl\{\log Z_{\lfloor n(1-\lambda)\mathbf{u}_2\rfloor} \ge n(1-\lambda)r_2\bigr\} + o(1) \end{aligned}$$

and letting $n \to \infty$ gives

$$J_{\mathbf{u}}(r) \le \lambda J_{\mathbf{u}_1}(r_1) + (1-\lambda) J_{\mathbf{u}_2}(r_2). \tag{3.14}$$

Finiteness of J follows from (3.5), so now we know J to be a finite, convex function on $(\mathbb{R}_+^d \setminus \{\mathbf{0}\}) \times \mathbb{R}$. This implies that J is continuous in the interior of $(\mathbb{R}_+^d \setminus \{\mathbf{0}\}) \times \mathbb{R}$ and upper semicontinuous on the whole set $(\mathbb{R}_+^d \setminus \{\mathbf{0}\}) \times \mathbb{R}$ [28], Theorems 10.1 and 10.2.

The law of large numbers for the free energy implies $J_{\mathbf{u}}(r) = 0$ for $r < p(\mathbf{u})$ and then by continuity for $r \le p(\mathbf{u})$. With a minor adaptation of [9], Proposition 3.1(b), we get a concentration inequality: given $\mathbf{u}$, for $\varepsilon > 0$ there exists a constant $c > 0$ such that

$$\mathbb{P}\bigl\{|\log Z_{\lfloor n\mathbf{u}\rfloor} - \mathbb{E}\log Z_{\lfloor n\mathbf{u}\rfloor}| \ge n\varepsilon\bigr\} \le 2\exp\bigl(-c\varepsilon^2 n\bigr) \qquad \text{for all } n \in \mathbb{N}. \tag{3.15}$$

Since $n^{-1}\mathbb{E}\log Z_{\lfloor n\mathbf{u}\rfloor} \to p(\mathbf{u})$, this implies that $J_{\mathbf{u}}(r) > 0$ for $r > p(\mathbf{u})$.

We do a coupling proof for lower semicontinuity. Let $(\mathbf{u}, r) \to (\mathbf{v}, s)$ in $(\mathbb{R}_+^d \setminus \{\mathbf{0}\}) \times \mathbb{R}$. If each coordinate $v_i > 0$, then we have continuity $J_{\mathbf{u}}(r) \to J_{\mathbf{v}}(s)$ because convexity already gives continuity in the interior. Thus we may assume that some coordinates of $\mathbf{v}$ are zero. Since coordinates can be permuted without changing J, let us assume that $\mathbf{v} = (v_1, v_2, \dots, v_k, 0, \dots, 0)$ for a fixed $1 \le k < d$ where $v_1, \dots, v_k > 0$. If eventually $\mathbf{u}$ is also of the form $\mathbf{u} = (u_1, u_2, \dots, u_k, 0, \dots, 0)$ for the same k, then we are done by convexity-implied continuity again, this time in the interior of $(\mathbb{R}_+^k \setminus \{\mathbf{0}\}) \times \mathbb{R}$.

The remaining case is the one where $u_1, \dots, u_k > 0$ and $(u_{k+1}, \dots, u_d) \to \mathbf{0}$. We develop a family of couplings that eliminates these $d-k$ last coordinates one by one, starting with u_d, and puts us back in the interior case with continuity. Denote a lower-dimensional projection by $\mathbf{u}_{1,k} = (u_1, u_2, \dots, u_k)$.

The set of paths $\Pi_{\lfloor n\mathbf{u}\rfloor}$ is decomposed according to the locations of the $\lfloor nu_d\rfloor$ unit jumps in the $\mathbf{e}_d$-direction. The projections of these locations form a vector π from the set

$$\Lambda_{\lfloor n\mathbf{u}\rfloor} = \bigl\{\pi = \{\mathbf{x}^i\}_{i=0}^{\lfloor nu_d\rfloor+1} \in \bigl(\mathbb{Z}_+^{d-1}\bigr)^{\lfloor nu_d\rfloor+2} : \\ \mathbf{0} = \mathbf{x}^0 \le \mathbf{x}^1 \le \cdots \le \mathbf{x}^{\lfloor nu_d\rfloor+1} = \lfloor n\mathbf{u}_{1,d-1}\rfloor\bigr\}.$$

The partition function then decomposes according to the following jump locations:

$$Z_{\lfloor n\mathbf{u}\rfloor} = \sum_{\pi\in\Lambda_{\lfloor n\mathbf{u}\rfloor}} Z_{(\mathbf{0},0),(\mathbf{x}^1,0)} \prod_{i=1}^{\lfloor nu_d\rfloor} Z^{\square}_{(\mathbf{x}^i,i),(\mathbf{x}^{i+1},i)} \equiv \sum_{\pi\in\Lambda_{\lfloor n\mathbf{u}\rfloor}} Z_\pi, \tag{3.16}$$

where the last equality defines the $d-1$-dimensional partition functions Z_π.

For a fixed π, define a new environment $\widetilde{\omega}$ indexed by $\mathbb{Z}_+^{d-1}$ with this recipe:

(i) For $0 \le i \le \lfloor nu_d\rfloor$: for $\mathbf{y} \in \mathbb{Z}_+^{d-1}$ such that $\mathbf{x}^i \le \mathbf{y} \le \mathbf{x}^{i+1}$ but $\mathbf{y} \ne \mathbf{x}^i$, set $\widetilde{\omega}(\mathbf{y}) = \omega(\mathbf{y}, i)$.

(ii) $\widetilde{\omega}(\mathbf{0}) = \omega(\mathbf{0}, 0)$ and for $1 \le i \le \lfloor nu_d\rfloor$, $\widetilde{\omega}(\lfloor n\mathbf{u}_{1,d-1}\rfloor + i\mathbf{e}_{d-1}) = \omega(\mathbf{x}^i, i)$.

(iii) Pick all other $\widetilde{\omega}(\mathbf{y})$ independently of everything else.

Now, keeping π fixed, we project the paths down to $\mathbb{Z}_+^{d-1}$ and create a partition function (marked by a tilde) in the new environment $\widetilde{\omega}$:

$$\begin{aligned}\log Z_\pi &= \log Z_{(\mathbf{0},0),(\mathbf{x}^1,0)} + \sum_{i=1}^{\lfloor nu_d\rfloor} \log Z^{\square}_{(\mathbf{x}^i,i),(\mathbf{x}^{i+1},i)} \\ &= \sum_{i=0}^{\lfloor nu_d\rfloor} \log Z_{(\mathbf{x}^i,i),(\mathbf{x}^{i+1},i)} + \sum_{i=1}^{\lfloor nu_d\rfloor} \omega\bigl(\mathbf{x}^i, i\bigr) \\ &= \sum_{i=0}^{\lfloor nu_d\rfloor} \log \widetilde{Z}_{\mathbf{x}^i,\mathbf{x}^{i+1}} + \sum_{i=1}^{\lfloor nu_d\rfloor} \widetilde{\omega}\bigl(\lfloor n\mathbf{u}_{1,d-1}\rfloor + i\mathbf{e}_{d-1}\bigr) \\ &\le \log \widetilde{Z}_{\lfloor n\mathbf{u}_{1,d-1}\rfloor + \lfloor nu_d\mathbf{e}_{d-1}\rfloor}.\end{aligned} \tag{3.17}$$

Introduce the continuous functions ($1 \le i < d$)

$$F_i(\mathbf{u}) = \sum_{j=1}^{i-1} \bigl((u_j + u_i)\log(u_j + u_i) - u_j \log u_j - u_i \log u_i\bigr). \tag{3.18}$$

Counting the number of ways to decompose the length from 0 to $\lfloor nu_i \rfloor$ into $\lfloor nu_d \rfloor + 1$ segments and Stirling's formula give

$$\begin{aligned} m_0 = |\Lambda_{\lfloor n\mathbf{u} \rfloor}| &= \prod_{1 \le i \le d-1} \binom{\lfloor nu_i \rfloor + \lfloor nu_d \rfloor}{\lfloor nu_d \rfloor + 1} = \exp\{nF_d(\mathbf{u}) + o(n)\} \\ &\le \exp\{nF_d(\mathbf{u}) + n\eta\}, \end{aligned} \tag{3.19}$$

where the last inequality is valid for large n and we introduced a small $\eta > 0$ that we can send to zero after limits in n have been taken. By a union bound and the coupling (3.17) separately for each $\pi \in \Lambda_{\lfloor n\mathbf{u} \rfloor}$,

$$\begin{aligned} -J_{\mathbf{u}}(r) &\le \varlimsup_{n\to\infty} n^{-1} \log \sum_{\pi \in \Lambda_{\lfloor n\mathbf{u} \rfloor}} \mathbb{P}\{\log Z_\pi \ge nr - \log m_0\} \\ &\le \lim_{n\to\infty} \biggl(\frac{\log m_0}{n} \\ &\qquad + n^{-1} \log \mathbb{P}\{\log \widetilde{Z}_{\lfloor n\mathbf{u}_{1,d-1} \rfloor + \lfloor nu_d \mathbf{e}_{d-1} \rfloor} \ge nr - nF_d(\mathbf{u}) - n\eta\} \biggr) \\ &= F_d(\mathbf{u}) - J_{\mathbf{u}_{1,d-1} + u_d \mathbf{e}_{d-1}}\bigl(r - F_d(\mathbf{u}) - \eta\bigr). \end{aligned}$$

In the last step above a little correction as in (3.12) replaces $\lfloor n\mathbf{u}_{1,d-1} \rfloor + \lfloor nu_d \mathbf{e}_{d-1} \rfloor$ with $\lfloor n\mathbf{u}_{1,d-1} + nu_d \mathbf{e}_{d-1} \rfloor$.

Let $\widetilde{\mathbf{u}}_{1,d} = \mathbf{u}$ and for $1 \le i < d$,

$$\widetilde{\mathbf{u}}_{1,i} = \mathbf{u}_{1,i} + \sum_{j=i+1}^{d} u_j \mathbf{e}_i \in \mathbb{Z}_+^i.$$

Proceeding inductively, we get the lower bound

$$J_{\mathbf{u}}(r) \ge J_{\widetilde{\mathbf{u}}_{1,k}}\biggl(r - \sum_{k+1 \le i \le d} \bigl(F_i(\mathbf{u}) - \eta\bigr)\biggr) - \sum_{k+1 \le i \le d} F_i(\mathbf{u}). \tag{3.20}$$

On the right-hand side we have a rate function $J_{\widetilde{\mathbf{u}}_{1,k}}$ with $\widetilde{\mathbf{u}}_{1,k} \to \mathbf{v}_{1,k}$ in the interior of $\mathbb{R}_+^k$. Thus we have continuity. We can first let $\eta \searrow 0$. Then let $(\mathbf{u}, r) \to (\mathbf{v}, s)$. Note that $u_i \to 0$ implies $F_i(\mathbf{u}) \to 0$. Together all this gives the lower semicontinuity

$$\varliminf_{(\mathbf{u},r)\to(\mathbf{v},s)} J_{\mathbf{u}}(r) \ge J_{\widetilde{\mathbf{v}}_{1,k}}(s) = J_{\mathbf{v}}(s).$$

Now we know J is continuous on all of $(\mathbb{R}_+^d \setminus \{\mathbf{0}\}) \times \mathbb{R}$.

Let us observe limit (3.7). From one side we have

$$\mathbb{P}\{\log Z^{\square}_{\lfloor n\mathbf{u}\rfloor} \ge nr\} \ge \mathbb{P}\{\log Z_{\lfloor n\mathbf{u}\rfloor} \ge nr\}\mathbb{P}\{\omega(\mathbf{0}) \ge 0\}.$$

From the other, pick a coordinate $u_i > 0$, and for each n an integer $n < m_n < n + o(n)$ such that $2\mathbf{e}_i + \lfloor n\mathbf{u}\rfloor \le \lfloor m_n\mathbf{u}\rfloor$. For each n fix a directed path $\{x^n_j\}$ from $2\mathbf{e}_i + \lfloor n\mathbf{u}\rfloor$ to $\lfloor m_n\mathbf{u}\rfloor$. Inequality

$$\omega(\mathbf{e}_i) + \log Z^{\square}_{2\mathbf{e}_i, 2\mathbf{e}_i+\lfloor n\mathbf{u}\rfloor} + \sum_j \omega(x^n_j) \le \log Z_{\lfloor m_n\mathbf{u}\rfloor}$$

gives

$$\mathbb{P}\{\log Z^{\square}_{\lfloor n\mathbf{u}\rfloor} \ge nr\}\mathbb{P}\Big\{\omega(\mathbf{e}_i) + \sum_j \omega(x^n_j) \ge 0\Big\} \le \mathbb{P}\{\log Z_{\lfloor m_n\mathbf{u}\rfloor} \ge nr\}.$$

Assumption (3.5) and the continuity of J give the conclusion. $\square$

PROOF OF THEOREM 3.3. It remains to prove continuity at $(\mathbf{0}, s)$. Let $(\mathbf{u}, r) \to (\mathbf{0}, s)$. Define the right-tail Cramér rate function for $a > 0$, $x \in \mathbb{R}$:

$$\begin{aligned}\kappa_a(x) &= -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\Big\{n^{-1}\sum_{i=1}^{\lfloor na\rfloor} \omega(x_i) \ge nx\Big\}\\ &= \begin{cases} aI(x/a), & x \ge a\mathbb{E}[\omega(\mathbf{0})],\\ 0, & x \le a\mathbb{E}[\omega(\mathbf{0})].\end{cases}\end{aligned}$$

Check that as $(a, x) \to (0, s)$, $\kappa_a(x) \to J_{\mathbf{0}}(s)$ defined by (3.9).

For upper semicontinuity, bound $Z_{\lfloor n\mathbf{u}\rfloor}$ below by a single path

$$J_{\mathbf{u}}(r) \le -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\Big\{n^{-1}\sum_{i=1}^{|\lfloor n\mathbf{u}\rfloor|_1} \omega(x_i) \ge nr\Big\} = \kappa_{|\mathbf{u}|_1}(r).$$

For lower semicontinuity, permute the coordinates so that $u_1 > 0$ as $\mathbf{u} \to \mathbf{0}$. Apply (3.20) after η has been taken to zero:

$$J_{\mathbf{u}}(r) \ge J_{u_1\mathbf{e}_1}\Big(r - \sum_{2\le i\le d} F_i(\mathbf{u})\Big) - \sum_{2\le i\le d} F_i(\mathbf{u}).$$

Since $J_{u_1\mathbf{e}_1} = \kappa_{u_1}$ we get the lower semicontinuity. $\square$

Finally two lemmas for later use. The next one allows more general lattice sequences for the right-tail LDP.

LEMMA 3.5. *Let $\mathbf{y} \in (0, \infty)^d$ and $\mathbf{u}_n \in \mathbb{Z}^d_+$ be a sequence such that $n^{-1}\mathbf{u}_n \to \mathbf{y}$. Then for $r \in \mathbb{R}$,*

$$\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\mathbf{u}_n} \ge nr\} = -J_{\mathbf{y}}(r). \tag{3.21}$$

PROOF. Let us use assumption (3.5) again. Since the coordinates of $\mathbf{u}_n$ and $\lfloor n\mathbf{y} \rfloor$ are increasing to ∞, for each n we can find ℓ_n and m_n such that $\lfloor \ell_n \mathbf{y} \rfloor \le \mathbf{u}_n \le \lfloor m_n \mathbf{y} \rfloor$ and in such a way that $n - \ell_n$, $n - m_n$ are eventually $o(n)$. For each n fix directed paths $\{x_{n,i}\}_{0 \le i \le K_n}$ from $\lfloor \ell_n \mathbf{y} \rfloor$ to $\mathbf{u}_n$ and $\{x'_{n,j}\}_{0 \le j \le K'_n}$ from $\mathbf{u}_n$ to $\lfloor m_n \mathbf{y} \rfloor$. Then

$$Z_{\lfloor \ell_n \mathbf{y} \rfloor} \cdot W_n \le Z_{\mathbf{u}_n} \le Z_{\lfloor m_n \mathbf{y} \rfloor} \cdot \bigl(W'_n\bigr)^{-1},$$

where

$$\log W_n = \sum_{1 \le i \le K_n} \omega(x_i) \quad \text{and} \quad \log W'_n = \sum_{1 \le i \le K'_n} \omega\bigl(x'_i\bigr).$$

Assumption $n^{-1}\mathbf{u}_n \to \mathbf{y}$ implies that K_n and K'_n are also $o(n)$.

The estimates we need follow. For example,

$$\mathbb{P}\{\log Z_{\lfloor m_n \mathbf{y} \rfloor} \ge nr\} \ge \mathbb{P}\bigl\{\log W'_n \ge 0\bigr\} \mathbb{P}\{\log Z_{\mathbf{u}_n} \ge nr\}$$

and then by assumption (3.5) and the continuity of the rate function,

$$\varlimsup_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\mathbf{u}_n} \ge nr\} \le \lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\lfloor m_n \mathbf{y} \rfloor} \ge nr\} = -J_{\mathbf{y}}(r).$$

Similarly for the complementary lower bound on $\varliminf$. $\square$

LEMMA 3.6. *Suppose that for each n, L_n and Z_n are independent random variables. Assume that the limits*

$$\lambda(s) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{L_n \ge ns\}, \tag{3.22}$$

$$\phi(s) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{Z_n \ge ns\} \tag{3.23}$$

exist and are finite for all $s \in \mathbb{R}$. Assume that $\lambda(a_\lambda) = \phi(a_\phi) = 0$ for some a_λ, $a_\phi \in \mathbb{R}$. Assume also that λ is continuous. Then for $r \in \mathbb{R}$

$$\begin{aligned} &\lim_{n\to\infty} \frac{\log \mathbb{P}\{L_n + Z_n \ge nr\}}{n} \\ &\quad = \begin{cases} -\inf_{a_\lambda \le s \le r - a_\phi} \bigl\{\phi(r-s) + \lambda(s)\bigr\}, & r > a_\phi + a_\lambda, \\ 0, & r \le a_\phi + a_\lambda. \end{cases} \end{aligned} \tag{3.24}$$

PROOF. The lower bound $\ge$ follows from

$$\mathbb{P}\{L_n + Z_n \ge nr\} \ge \mathbb{P}\{L_n \ge ns\} \mathbb{P}\bigl\{Z_n \ge n(r-s)\bigr\}.$$

Since an upper bound 0 is obvious, it remains to show the upper bound for the case $r > a_\phi + a_\lambda$. Take a finite partition $a_\lambda = q_0 < \cdots < q_m = r - a_\phi$. Then use a union bound and independence:

$$
\begin{aligned}
&\mathbb{P}\{L_n + Z_n \ge nr\}\\
&\quad \le \mathbb{P}\{L_n + Z_n \ge nr, L_n < nq_0\}\\
&\qquad + \sum_{i=0}^{m-1} \mathbb{P}\{L_n + Z_n \ge nr, nq_i \le L_n \le nq_{i+1}\} + \mathbb{P}\{L_n \ge nq_m\}\\
&\quad \le \mathbb{P}\big\{Z_n \ge n(r-q_0)\big\} + \sum_{i=0}^{m-1} \mathbb{P}\big\{Z_n \ge n(r-q_{i+1})\big\}\mathbb{P}\{L_n \ge nq_i\}\\
&\qquad + \mathbb{P}\{L_n \ge nq_m\}.
\end{aligned}
$$

From this,

$$
\begin{aligned}
&\overline{\lim_{n\to\infty}}\, n^{-1} \log \mathbb{P}\{L_n + Z_n \ge nr\}\\
&\quad \le -\min\Big\{\phi(r-q_0), \min_{0\le i\le m-1}\big[\phi(r-q_{i+1}) + \lambda(q_i)\big], \lambda(q_m)\Big\}.
\end{aligned}
$$

Note that $\lambda(q_0) = \phi(r - q_m) = 0$, refine the partition and use the continuity of λ. $\square$

4. Proofs for the i.i.d. log-gamma model. In this section we prove the results of Section 2.1. Throughout this section the dimension $d = 2$ and the weights satisfy $Y_{i,j}^{-1} \sim \text{Gamma}(\mu)$ as in (2.1). As before, for $(s,t) \in \mathbb{R}_+^2 \setminus \{(0,0)\}$ define the function $J_{s,t}$ by the limit

$$
J_{s,t}(r) = -\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor} \ge nr\}, \qquad r \in \mathbb{R}. \tag{4.1}
$$

At the origin set

$$
J_{0,0}(r) = \begin{cases} 0, & r \le 0,\\ \mu r, & r \ge 0. \end{cases} \tag{4.2}
$$

Then, as observed in Remark 3.4, the function $J_{s,t}(r)$ is finite and continuous at all $(s,t,r) \in \mathbb{R}_+^2 \times \mathbb{R}$.

We begin with a lemma that proves Theorem 2.1.

LEMMA 4.1. *For $(s,t) \in \mathbb{R}_+^2$ the limiting free energy of (2.5) satisfies*

$$
p(s,t) = \inf_{0<\theta<\mu}\big\{-s\Psi_0(\theta) - t\Psi_0(\mu-\theta)\big\}. \tag{4.3}
$$

The infimum is achieved at some θ because $\Psi_0(0+) = -\infty$.

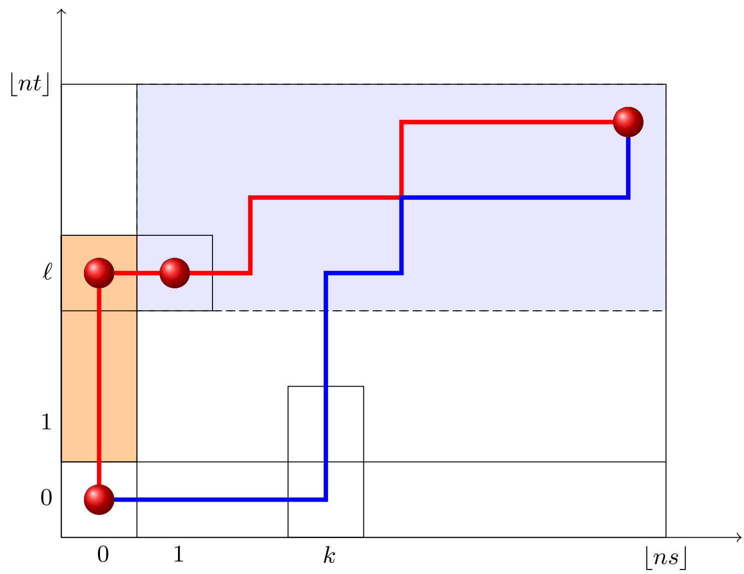

FIG. 2. *Graphical representation of the decomposition in equation* (4.4).

PROOF. The proof anticipates some themes of the later LDP proof, but in a simpler context. We already recorded the law of large numbers (2.18). The decomposition (see Figure 2)

$$
\begin{aligned}
Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} &= \sum_{k=1}^{\lfloor ns\rfloor}\Bigl(\prod_{i=1}^{k} U_{i,0}\Bigr) Z^{\square}_{(k,1),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \\
&\quad + \sum_{\ell=1}^{\lfloor nt\rfloor}\Bigl(\prod_{j=1}^{\ell} V_{0,j}\Bigr) Z^{\square}_{(1,\ell),(\lfloor ns\rfloor,\lfloor nt\rfloor)}
\end{aligned}
\tag{4.4}
$$

from (2.14) gives asymptotically

$$
\begin{aligned}
&\lim_{n\to\infty} n^{-1}\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\
&\quad = \lim_{n\to\infty}\Biggl\{\max_{1\le k\le\lfloor ns\rfloor}\Biggl(n^{-1}\sum_{i=1}^{k}\log U_{i,0} + n^{-1}\log Z^{\square}_{(k,1),(\lfloor ns\rfloor,\lfloor nt\rfloor)}\Biggr) \\
&\qquad\qquad \vee \max_{1\le \ell\le\lfloor nt\rfloor}\Biggl(n^{-1}\sum_{j=1}^{\ell}\log V_{0,j} + n^{-1}\log Z^{\square}_{(1,\ell),(\lfloor ns\rfloor,\lfloor nt\rfloor)}\Biggr)\Biggr\}.
\end{aligned}
$$

This can be coarse-grained with readily controllable errors of sums of independent variables. We omit the details since similar arguments appear elsewhere in the paper. The conclusion is the alternative formula

$$
\begin{aligned}
&\lim_{n\to\infty} n^{-1}\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\\
&\quad = \sup_{0\le a\le s}\bigl\{-a\Psi_0(\theta)+p(s-a,t)\bigr\}\\
&\qquad \vee \sup_{0\le b\le t}\bigl\{-b\Psi_0(\mu-\theta)+p(s,t-b)\bigr\}.
\end{aligned}
\tag{4.5}
$$

Take $s=t$, combine (2.18) and (4.5), and use the symmetry $p(s,t)=p(t,s)$ to get

$$-t\bigl(\Psi_0(\theta)+\Psi_0(\mu-\theta)\bigr)=\sup_{0\le a\le t}\bigl\{-a\bigl(\Psi_0(\theta)\wedge\Psi_0(\mu-\theta)\bigr)+p(t-a,t)\bigr\}.$$

Take $\theta\in(0,\mu/2]$ so that $\Psi_0(\theta)\le\Psi_0(\mu-\theta)$ (Ψ_0 is strictly increasing) and set $a=t-s$:

$$-t\Psi_0(\mu-\theta)=\sup_{0\le s\le t}\bigl\{s\Psi_0(\theta)+p(s,t)\bigr\}.$$

Turn this into a convex duality through the change of variable $v=\Psi_0(\theta)$:

$$-t\Psi_0\bigl(\mu-\Psi_0^{-1}(v)\bigr)=\sup_{0\le s\le t}\bigl\{sv+p(s,t)\bigr\},\qquad v\in\bigl(-\infty,\Psi_0(\mu/2)\bigr]. \tag{4.6}$$

It follows from the limit definition of $p(s,t)$ that it is concave and continuous in $s\in[0,t]$. Extend $f(s)=-p(s,t)$ to a lower semicontinuous convex function of $s\in\mathbb{R}$ by setting $f(s)=\infty$ for $s\notin[0,t]$. Then (4.6) tells us that

$$f^*(v)=-t\Psi_0\bigl(\mu-\Psi_0^{-1}(v)\bigr)\qquad\text{for } v\in\bigl(-\infty,\Psi_0(\mu/2)\bigr].$$

We can differentiate to get $\lim_{v\searrow-\infty}(f^*)'(v)=0$ and $(f^*)'(\Psi_0(\mu/2))=t$. These derivative values imply that for $s\in[0,t]$, the supremum in the double convex duality can be restricted as follows:

$$f(s)=\sup_{v\in(-\infty,\Psi_0(\mu/2)]}\bigl\{vs-f^*(v)\bigr\}.$$

Undoing the change of variables turns this equation into (4.3) which is thereby proved. $\square$

The next lemma gives left tail bounds strong enough to imply $I_{s,t}(r)=\infty$ for $r<p(s,t)$, and the same result for the stationary model. The proof is a straightforward coarse-graining argument. We do not expect the results to be optimal.

LEMMA 4.2. *Fix $0<a<1$. Then there exist constants $0<c,C<\infty$ that depend on the parameters given below, so that the following estimates hold*:

(a) *For $(s,t)\in(0,\infty)^2$ and $r<p(s,t)$*,

$$\mathbb{P}\{\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor}\le nr\}\le Ce^{-cn^{1+a}}\qquad\textit{for all } n\ge 1. \tag{4.7}$$

(b) *For* $(s,t)=\alpha(\Psi_1(\mu-\theta),\Psi_1(\theta))$ *for some* $\alpha>0$, *parallel to the characteristic direction*, *and* $r<p^{(\theta)}(s,t)$,

$$\mathbb{P}\{\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\le nr\}\le Ce^{-cn^{1+a}}\qquad \textit{for all } n\ge 1. \tag{4.8}$$

PROOF. We give a proof of (b) with some details left sketchy. Part (a) has a similar proof. We bound $Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ from below by considering a subset of lattice paths, arranged in a collection of i.i.d. partition functions over subsets of the rectangle.

The choice of (s,t) implies that $p^{(\theta)}(s,t)=p(s,t)$. Fix $0<\varepsilon<(p^{(\theta)}(s,t)-r)/4$. Fix $m\in\mathbb{N}$ large enough so that $m(s\wedge t)\ge 1$ and

$$\mathbb{E}\log Z_{\lfloor ms\rfloor,\lfloor mt\rfloor}>m(r+2\varepsilon). \tag{4.9}$$

Let $B^{k,\ell}_{a,b}=\{a,\dots,a+k-1\}\times\{b,\dots,\ell+b-1\}$ denote the $k\times\ell$ rectangle with lower left corner at (a,b). For $i,\ell\ge 0$ define pairwise disjoint $\lfloor ms\rfloor\times\lfloor mt\rfloor$ rectangles

$$B^i_\ell=B^{\lfloor ms\rfloor,\lfloor mt\rfloor}_{(\ell+i)\lfloor ms\rfloor-\ell+1,\ell\lfloor mt\rfloor+1}.$$

Define a diagonal union of these rectangles by $\Delta_i=\bigcup_{\ell\ge 0}B^i_\ell$, $i\ge 0$; see Figure 3.

Let $M=\lfloor n^a\rfloor\lfloor ms\rfloor$. This is the range of diagonals Δ_i we consider. Then we cut the diagonals off before they exit the $\lfloor ns\rfloor\times\lfloor nt\rfloor$ rectangle. Let $N=N(n)$ be the maximal integer such that B^M_N lies in $[0,\lfloor ns\rfloor]\times[0,\lfloor nt\rfloor]$. Diagonal Δ_M exits the

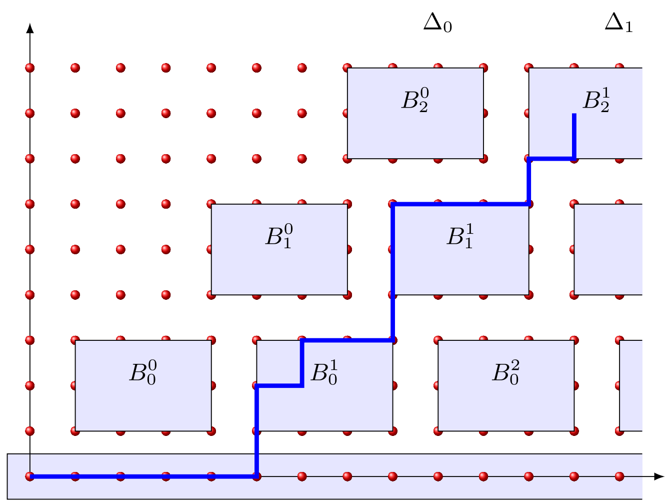

FIG. 3. *The* $\lfloor ms\rfloor\times\lfloor mt\rfloor$ *rectangles and the diagonals* Δ_i *in the proof of Lemma* 4.2. *The thickset line is a lattice path that is counted in* Z_1.

$\lfloor ns \rfloor \times \lfloor nt \rfloor$ rectangle through the east edge, and consequently there exist positive constants c_m, C_m such that

$$\begin{aligned} &\lfloor ns \rfloor - c_m < N\lfloor ms \rfloor + \lfloor ms \rfloor \lfloor n^a \rfloor \le \lfloor ns \rfloor \quad \text{and} \\ &\lfloor nt \rfloor - C_m n^a < N\lfloor mt \rfloor \le \lfloor nt \rfloor. \end{aligned} \tag{4.10}$$

Having defined the cutoff N, define the remaining diagonals by $\Delta_i^n = \bigcup_{0 \le \ell \le N} B_\ell^i$ for $0 \le i \le M$. These diagonals lie in $[0, \lfloor ns \rfloor] \times [0, \lfloor nt \rfloor]$. Fix a path π that proceeds horizontally from point $(N\lfloor ms \rfloor, N\lfloor mt \rfloor + 1)$ to $(\lfloor ns \rfloor, N\lfloor mt \rfloor + 1)$ and then vertically up to $(\lfloor ns \rfloor, \lfloor nt \rfloor)$. The number of lattice points on π is a constant multiple of n^a.

For $0 \le i \le M$, let Z_i denote the partition function of paths $x_\cdot$ of the following type: $x_\cdot$ proceeds along the x-axis from the origin to $(i\lfloor ms \rfloor + 1, 0)$, enters Δ_i^n at $(i\lfloor ms \rfloor + 1, 1)$, and stays in Δ_i^n until it exits from the upper right corner of B_N^i with a vertical step that connects it with π. After that $x_\cdot$ follows π to $(\lfloor ns \rfloor, \lfloor nt \rfloor)$. The number K of points on $x_\cdot$ outside Δ_i^n is independent of i and bounded by a constant multiple of n^a. Let

$$X = \min\{Y_x : x \in \pi \text{ or } x \in \{(i, 0) : 0 \le i \le M\}\}$$

be the minimal weight outside Δ_i^n encountered by any path $x_\cdot$ of Z_i, for any $0 \le i \le M$.

Let Z_i^Δ be the partition function of all lattice paths in Δ_i^n from the lower left corner of B_0^i to the upper right corner of B_N^i. Then $Z_i \ge X^K Z_i^\Delta$, and consequently

$$\begin{aligned} \mathbb{P}\{\log Z^{(\theta)}_{\lfloor ns \rfloor, \lfloor nt \rfloor} \le nr\} &\le \mathbb{P}\left\{\log \sum_{i=0}^{M} X^K Z_i^\Delta \le nr\right\} \\ &= \mathbb{P}\left\{K \log X + \log \sum_{i=0}^{M} Z_i^\Delta \le nr\right\} \\ &\le \mathbb{P}\{K \log X \le -n\varepsilon\} \\ &\quad + \mathbb{P}\left\{\log \sum_{i=0}^{M} Z_i^\Delta \le n(r + \varepsilon)\right\}. \end{aligned} \tag{4.11}$$

Explicit computation with the gamma distribution and $K \le cn^a$ give the probability $\mathbb{P}\{K \log X \le -n\varepsilon\} \le e^{-n^2}$ for large n.

The $\{Z_i^\Delta\}$ are i.i.d., and Z_0^Δ is a product of the i.i.d. partition functions Z_k^0 of the individual rectangles B_k^0 whose mean was controlled by (4.9). A standard large deviation estimate for an i.i.d. sum gives

$$\begin{aligned}
\mathbb{P}\left\{\log \sum_{i=0}^{M} Z_i^{\Delta} \le n(r+\varepsilon)\right\} &\le \mathbb{P}\{\log Z_0^{\Delta} \le n(r+\varepsilon)\}^M \\
&= \mathbb{P}\left\{\sum_{k=0}^{N} \log Z_k^0 \le n(r+\varepsilon)\right\}^M \\
&= \mathbb{P}\left\{\sum_{k=0}^{n/m+o(n)} \log Z_k^0 \le n(r+\varepsilon)\right\}^M \\
&\le e^{-cnM} \le e^{-c_1 n^{1+a}}.
\end{aligned}$$

Putting these bounds back on line (4.11) completes the proof of (4.8). $\square$

The main work resides in proving the following right tail result.

PROPOSITION 4.3. *Let $(s,t) \in \mathbb{R}^2$. Then for all $r \in \mathbb{R}$, $J_{s,t}(r)$ is given by*

$$J_{s,t}(r) = \sup_{\xi \in [0,\mu)} \Big\{ r\xi - \inf_{\theta \in (\xi,\mu)} \big(tM_\theta(\xi) - sM_{\mu-\theta}(-\xi)\big)\Big\}. \tag{4.12}$$

Before turning to the proof of Proposition 4.3 let us observe how Theorem 2.2 follows.

PROOF OF THEOREM 2.2. Only a few simple observations are required. Start by defining $I_{s,t}$ as given in (2.6). Then formula (2.9) that connects $I_{s,t}$ and $J_{s,t}$ is established by (4.12) and by knowing that $J_{s,t}(r) = 0$ for $r \le p(s,t)$ (Theorem 3.2). The regularity properties of $I_{s,t}$ follow from the general properties of J in Theorems 3.2 and 3.3.

The upper large deviation bound (2.8) is built into (4.7) and (4.1).

For the lower large deviation bound (2.7), we consider three cases:

(i) If $p(s,t) \in G$, then $\mathbb{P}\{n^{-1} \log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor} \in G\} \to 1$ and (2.7) holds trivially because its right-hand side is ≤ 0.

(ii) If $G \subseteq (-\infty, p(s,t))$, (2.7) holds trivially because its right-hand side is $-\infty$.

(iii) The remaining case is the one where G contains an interval $(a,b) \subset (p(s,t),\infty)$. Since the distribution is continuous including a into G makes no difference, and so

$$\begin{aligned}
&n^{-1}\log \mathbb{P}\{n^{-1}\log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor} \in G\} \\
&\quad \ge n^{-1}\log\big(\mathbb{P}\{\log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor} \ge na\} - \mathbb{P}\{\log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor} \ge nb\}\big) \\
&\quad \longrightarrow -J_{s,t}(a),
\end{aligned}$$

where the limit follows from (4.1) and the strict increasingness of $J_{s,t}$ on $[p(s,t),\infty)$ which implies that for large enough n,

$$\mathbb{P}\{\log Z_{\lfloor ns\rfloor,\lfloor nt\rfloor} \ge nb\} \le e^{-nJ_{s,t}(a)-n\varepsilon}$$

for some $\varepsilon > 0$. We can take $a = \inf G \cap (p(s,t),\infty)$ and then $J_{s,t}(a) = \inf_{r\in G\cap(p(s,t),\infty)} I_{s,t}(r) = \inf_{r\in G} I_{s,t}(r)$. $\square$

The remainder of the section is devoted to proving Proposition 4.3. Again we begin with the decomposition (4.4) of the stationary partition function. Inside the sums on the right-hand side of (4.4) we have partition functions with i.i.d. $\text{Gamma}^{-1}(\mu)$-weights $\{Y_{i,j}\}$ whose large deviations we wish to extract. But we do not know the large deviations of $\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$, so at first the decomposition seems unhelpful. To get around the problem, use definition (2.15) to write

$$\log Z^{(\theta)}_{\lfloor ns\rfloor,\lfloor nt\rfloor} - \log Z^{(\theta)}_{0,\lfloor nt\rfloor} = \sum_{j=1}^{\lfloor ns\rfloor} \log U_{i,\lfloor nt\rfloor}.$$

By Proposition 2.10 we have a sum of i.i.d.'s on the right, whose large deviations we can immediately write down by Cramér's theorem. To take advantage of this, divide through (4.4) by $Z^{(\theta)}_{0,\lfloor nt\rfloor} = \prod_{j=1}^{\lfloor nt\rfloor} V_{0,j}$ to rewrite it as

$$\begin{aligned}\prod_{i=1}^{\lfloor ns\rfloor} U_{i,\lfloor nt\rfloor} &= \sum_{\ell=1}^{\lfloor nt\rfloor}\Bigg(\prod_{j=\ell+1}^{\lfloor nt\rfloor} V_{0,j}^{-1}\Bigg) Z^{\square}_{(1,\ell),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \\ &\quad + \sum_{k=1}^{\lfloor ns\rfloor}\Bigg(\prod_{j=1}^{\lfloor nt\rfloor} V_{0,j}^{-1}\Bigg)\Bigg(\prod_{i=1}^{k} U_{i,0}\Bigg) Z^{\square}_{(k,1),(\lfloor ns\rfloor,\lfloor nt\rfloor)}.\end{aligned} \tag{4.13}$$

To compactify notation we use a convention where the y-axis is labeled by negative indices and introduce these quantities:

$$\text{for } k\in\mathbb{Z} \qquad \eta_k = \begin{cases} \displaystyle\prod_{j=-k+1}^{\lfloor nt\rfloor} V_{0,j}^{-1}, & k\le 0,\\ \displaystyle\Bigg(\prod_{j=1}^{\lfloor nt\rfloor} V_{0,j}^{-1}\Bigg)\prod_{i=1}^{k} U_{i,0}, & k\ge 1,\end{cases} \tag{4.14}$$

where an empty product equals 1 by definition, and

$$\text{for } z\in\mathbb{R} \qquad \mathbf{v}(z) = \begin{cases} (1,\lfloor -z\rfloor), & z\le -1,\\ (1,1), & -1<z<1,\\ (\lfloor z\rfloor,1), & z\ge 1.\end{cases} \tag{4.15}$$

Then (4.13) rewrites as

$$\prod_{i=1}^{\lfloor ns\rfloor} U_{i,\lfloor nt\rfloor} = \sum_{\substack{k=-\lfloor nt\rfloor\\ k\neq 0}}^{\lfloor ns\rfloor} \eta_k Z^{\square}_{\mathbf{v}(k),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \tag{4.16}$$

from which we extract these inequalities:

$$\begin{aligned}
&\log\eta_k + \log Z^{\square}_{\mathbf{v}(k),(\lfloor ns\rfloor,\lfloor nt\rfloor)}\\
&\quad\le \sum_{i=1}^{\lfloor ns\rfloor}\log U_{i,\lfloor nt\rfloor}\\
&\quad\le \max_{\substack{-\lfloor nt\rfloor\le k\le\lfloor ns\rfloor\\ k\neq 0}}\bigl\{\log\eta_k + \log Z^{\square}_{\mathbf{v}(k),(\lfloor ns\rfloor,\lfloor nt\rfloor)}\bigr\} + \log\bigl(n(s+t)\bigr).
\end{aligned} \tag{4.17}$$

These inequalities will be the basis for proving Proposition 4.3.

We record the right tail rate functions for the random variables in (4.17).

For the i.i.d. weights $\{U_{i,\lfloor nt\rfloor}\}$ we have the right branch of the Cramér rate function

$$\begin{aligned}
R_s(r) &= -\lim_{n\to\infty} n^{-1}\log\mathbb{P}\Biggl\{\sum_{i=1}^{\lfloor ns\rfloor}\log U_{i,\lfloor nt\rfloor}\ge nr\Biggr\}\\
&= \begin{cases} sI_\theta(rs^{-1}), & r\ge -s\Psi_0(\theta),\\ 0, & r< -s\Psi_0(\theta).\end{cases}
\end{aligned} \tag{4.18}$$

The rate function I_θ defined by (4.18) is given by

$$I_\theta(r) = -r\Psi_0^{-1}(-r) - \log\Gamma\bigl(\Psi_0^{-1}(-r)\bigr) + \theta r + \log\Gamma(\theta), \qquad r\in\mathbb{R}. \tag{4.19}$$

The convex dual of R_s is given by

$$R_s^*(\xi) = \begin{cases} s\log\Gamma(\theta-\xi) - s\log\Gamma(\theta), & 0\le\xi<\theta,\\ \infty, & \xi<0 \text{ or } \xi\ge\theta,\end{cases} \tag{4.20}$$

and we emphasize that it can be finite only when $\theta>\xi\ge 0$.

For real $a\in[-t,s]$,

$$\kappa_a(r) = -\lim_{n\to\infty} n^{-1}\log\mathbb{P}\{\log\eta_{\lfloor na\rfloor}\ge nr\} \tag{4.21}$$

exists and is finite, convex and continuous in r. (For $a\le 0$ it is simply a Cramér rate function for an i.i.d. sum, and for $a>0$ we can use Lemma 3.6.) The convex dual is

$$\kappa_a^*(\xi) = \sup_{r\in\mathbb{R}}\{\xi r - \kappa_a(r)\}$$

(4.22)

$$= \begin{cases} (t+a)\big(\log\Gamma(\mu-\theta+\xi) - \log\Gamma(\mu-\theta)\big), \\ \qquad\qquad -t \le a \le 0, \xi \ge 0, \\ t\big(\log\Gamma(\mu-\theta+\xi) - \log\Gamma(\mu-\theta)\big) \\ \qquad + a\big(\log\Gamma(\theta-\xi) - \log\Gamma(\theta)\big), \\ \qquad\qquad 0 < a \le s, 0 \le \xi < \theta, \\ \infty, \qquad\quad \text{otherwise.} \end{cases}$$

The derivation of (4.22) is similar to that of (4.20) from (4.18). Note that there is a discontinuity in κ_a and κ_a^* as a passes through 0. The rightmost zero $m_{\kappa,a}$ of κ_a is the law of large numbers limit,

$$m_{\kappa,a} = \lim_{n\to\infty} \frac{\log \eta_{\lfloor na\rfloor}}{n} = \begin{cases} (t+a)\Psi_0(\mu-\theta), & -t \le a \le 0, \\ t\Psi_0(\mu-\theta) - a\Psi_0(\theta), & 0 < a \le s. \end{cases} \tag{4.23}$$

In contrast to the functions κ_a and κ_a^*, $m_{\kappa,a}$ is continuous at $a = 0$. Introduce the "macroscopic" version of (4.15): for real a,

$$n^{-1}\mathbf{v}(na) \to \bar{\mathbf{v}}(a) = \begin{cases} (0,-a), & -t \le a \le 0, \\ (a,0), & 0 \le a \le s. \end{cases} \tag{4.24}$$

With this notation we have, again for real $a \in [-t, s]$, for the partition functions that appear in (4.16), the following large deviations:

$$J_{(s,t)-\bar{\mathbf{v}}(a)}(r) = -\lim_{n\to\infty} n^{-1}\log\mathbb{P}\big\{\log Z^{\square}_{\mathbf{v}(na),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \ge nr\big\}. \tag{4.25}$$

We used Lemma 3.5 to take care of the small discrepancy between $(\lfloor ns\rfloor, \lfloor nt\rfloor) - \mathbf{v}(na)$ and $\lfloor n((s,t) - \bar{\mathbf{v}}(a))\rfloor$, unless $a = -t$ or $a = s$ when this is a case of i.i.d. large deviations, and therefore simpler.

Let $m_{\kappa,a}$ and $m_{J,b}$ be the rightmost zeroes of κ_a and $J_{(s,t)-\bar{\mathbf{v}}(b)}$, respectively. For $(a,b) \in [-t,s]^2$, let

$$H^{a,b}_{s,t}(r) = \lim_{n\to\infty} n^{-1}\log\mathbb{P}\big\{\log\eta_{\lfloor na\rfloor} + \log Z^{\square}_{\mathbf{v}(nb),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \ge nr\big\}$$

(4.26)

$$= \begin{cases} 0, & r < m_{\kappa,a} + m_{J,b}, \\ \inf\limits_{m_{\kappa,a}\le x\le r-m_{J,b}}\big\{\kappa_a(x) + J_{(s,t)-\bar{\mathbf{v}}(b)}(r-x)\big\}, \\ & r \ge m_{\kappa,a} + m_{J,b}. \end{cases}$$

The existence of $H^{a,b}_{s,t}(r)$ and the second equality follow from Lemma 3.6. We need some regularity:

LEMMA 4.4. *Fix $0 < s, t < \infty$ and a compact set $K \subseteq \mathbb{R}$. Then $H^{a,b}_{s,t}(r)$ is uniformly continuous as a function of $(b, r) \in [-t, s] \times K$, uniformly in $a \in$*

$[-t, s]$. *That is*,

$$\lim_{\delta \searrow 0} \sup_{\substack{a,b,b' \in [-t,s], r,x \in K: \\ |b-b'| \le \delta, |r-x| \le \delta}} \left| H_{s,t}^{a,b}(r) - H_{s,t}^{a,b'}(x) \right| = 0. \tag{4.27}$$

PROOF. This follows from the explicit formula in (4.26). First, we have the joint continuity $(b, r) \mapsto J_{(s,t)-\bar{\mathbf{v}}(b)}(r)$ from Theorem 3.3. Second, we argue that x in the infimum can be restricted to a single compact set simultaneously for $(a, b, r) \in [-t, s]^2 \times K$. That $m_{\kappa,a}$ is bounded is evident from (4.23). To show that the upper bound $r - m_{J,b}$ of x is bounded above, we need to show a lower bound on $m_{J,b} = p((s, t) - \mathbf{v}(b))$. A lower bound on the free energy is easy: by discarding all but a single path,

$$p\bigl((s,t) - \mathbf{v}(b)\bigr) = \lim_{n\to\infty} n^{-1} \log Z^{\square}_{\lfloor n((s,t)-\bar{\mathbf{v}}(b))\rfloor} \ge -\bigl(s + t - |b|\bigr)\Psi_0(\mu). \qquad \square$$

We abbreviate $H_{s,t}^{a}(r) = H_{s,t}^{a,a}(r)$.

The unknown rate functions $J_{s,t}$ are now inside (4.26), while the other rates R_s and κ_a we know explicitly. The next lemma is the counterpart of (4.17) in terms of rate functions.

LEMMA 4.5. *Let $s, t > 0$ and $r \in \mathbb{R}$. Then*

$$R_s(r) = \inf_{-t \le a \le s} H_{s,t}^{a}(r). \tag{4.28}$$

PROOF. For any $a \in [-t, s]$, by the first inequality of (4.17),

$$\begin{aligned} -R_s(r) &= \lim_{n\to\infty} n^{-1} \log \mathbb{P}\left\{ \sum_{i=1}^{\lfloor ns \rfloor} \log U_{i,\lfloor nt \rfloor} \ge nr \right\} \\ &\ge \lim_{n\to\infty} n^{-1} \log \mathbb{P}\bigl\{ \log \eta_{\lfloor na \rfloor} + \log Z^{\square}_{\mathbf{v}(na),(\lfloor ns \rfloor, \lfloor nt \rfloor)} \ge nr \bigr\} \\ &\ge -H_{s,t}^{a}(r). \end{aligned} \tag{4.29}$$

Supremum over $a \in [-t, s]$ on the right gives $\le$ in (4.28).

To get $\ge$ in (4.28) we use the second inequality of (4.17) together with a partitioning argument. Let $\varepsilon > 0$. Note this technical point about handling the errors of the partitioning. With $B, \delta > 0$, Chebyshev's inequality and the l.m.g.f. of (2.2) give the bound

$$\mathbb{P}\left\{ \sum_{i=1}^{\lfloor n\delta \rfloor} \log Y_{i,1} \le -n\varepsilon \right\} \le e^{-nB(\varepsilon - B^{-1}\delta \log(\Gamma(\mu+B)/\Gamma(\mu)))} \le e^{-B\varepsilon n/2}, \tag{4.30}$$

where the second inequality comes from choosing $\delta = \delta(\varepsilon, B)$ small enough. The right tail for $\log Y$ does not give such a bound with an arbitrarily large B. Consequently we arrange the errors so that they can be bounded as above.

Given $B > 0$, fix a small enough $\delta > 0$ and let $-t = a_0 < a_1 < \cdots < a_q = 0 < \cdots < a_m = s$ be a partition of the interval $-[t, s]$ so that $|a_{i+1} - a_i| < \delta$. We illustrate how a term with index k from the right-hand side of (4.17) is reduced to a term involving only partition points. Consider the case $a_i \geq 0$ and let $\lfloor na_i \rfloor \leq k \leq \lfloor na_{i+1} \rfloor$:

$$\begin{aligned}
&\mathbb{P}\{\log \eta_k + \log Z^{\square}_{\mathbf{v}(k),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \geq nr\} \\
&\quad \leq \mathbb{P}\Bigg\{\log \eta_{\lfloor na_{i+1}\rfloor} + \log Z^{\square}_{\mathbf{v}(na_i),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \\
&\qquad\qquad - \sum_{j=k+1}^{\lfloor na_{i+1}\rfloor} \log U_{j,0} - \sum_{j=\lfloor na_i\rfloor}^{k-1} \log Y_{j,1} \geq nr\Bigg\} \\
&\quad \leq \mathbb{P}\{\log \eta_{\lfloor na_{i+1}\rfloor} + \log Z^{\square}_{\mathbf{v}(na_i),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \geq n(r-\varepsilon)\} \qquad (4.31)\\
&\qquad + \mathbb{P}\Bigg\{- \sum_{j=k+1}^{\lfloor na_{i+1}\rfloor} \log U_{j,0} - \sum_{j=\lfloor na_i\rfloor}^{k-1} \log Y_{j,1} \geq n\varepsilon\Bigg\} \\
&\quad \leq \mathbb{P}\{\log \eta_{\lfloor na_{i+1}\rfloor} + \log Z^{\square}_{\mathbf{v}(na_i),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \geq n(r-\varepsilon)\} \\
&\qquad + e^{-B\varepsilon n/2}.
\end{aligned}$$

On the other hand, if $a_i < 0$ and $\lfloor -na_{i+1}\rfloor < -k \leq \lfloor -na_i \rfloor$, then we would develop as follows:

$$\begin{aligned}
&\log \eta_k + \log Z^{\square}_{\mathbf{v}(k),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \\
&\quad \leq \log \eta_{\lfloor na_i\rfloor} - \sum_{j=-k+1}^{-\lfloor na_i\rfloor} \log V_{0,j} + \log Z^{\square}_{\mathbf{v}(na_{i+1}),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \\
&\qquad - \sum_{j=\lfloor -na_{i+1}\rfloor \vee 1}^{-k-1} \log Y_{1,j}
\end{aligned}$$

and get the same bound as on line (4.31) but with a_i and a_{i+1} switched around.

Now for $\geq$ in (4.28). Assume n is large enough so that $n\varepsilon > \log(ns + nt)$. Starting from (4.17),

$$\begin{aligned}
&n^{-1} \log \mathbb{P}\Bigg\{\sum_{i=1}^{\lfloor ns\rfloor} \log U_{i,\lfloor nt\rfloor} \geq nr\Bigg\} \\
&\quad \leq \max_{\substack{-\lfloor nt\rfloor \leq k \leq \lfloor ns\rfloor \\ k\neq 0}} n^{-1} \log \mathbb{P}\{\log \eta_k + \log Z^{\square}_{\mathbf{v}(k),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \geq n(r-\varepsilon)\} \\
&\qquad + n^{-1}\log(ns+nt)
\end{aligned}$$

$$\leq \max_{0\leq i\leq q-1} n^{-1}\log\bigl(\mathbb{P}\{\log\eta_{\lfloor na_i\rfloor}+\log Z^{\square}_{\mathbf{v}(na_{i+1}),(\lfloor ns\rfloor,\lfloor nt\rfloor)}\geq n(r-2\varepsilon)\}$$
$$+e^{-B\varepsilon n/2}\bigr)$$
$$\vee \max_{q\leq i\leq m-1} n^{-1}\log\bigl(\mathbb{P}\{\log\eta_{\lfloor na_{i+1}\rfloor}+\log Z^{\square}_{\mathbf{v}(na_i),(\lfloor ns\rfloor,\lfloor nt\rfloor)}$$
$$\geq n(r-2\varepsilon)\}$$
$$+e^{-B\varepsilon n/2}\bigr)+\varepsilon.$$

Take $n\to\infty$ above to obtain

$$-R_s(r)\leq\Bigl\{\max_{0\leq i\leq q-1}\bigl(-H^{a_i,a_{i+1}}_{s,t}(r-2\varepsilon)\bigr)\vee(-B\varepsilon/2)\Bigr\}$$
$$\vee\Bigl\{\max_{q\leq i\leq m-1}\bigl(-H^{a_{i+1},a_i}_{s,t}(r-2\varepsilon)\bigr)\vee(-B\varepsilon/2)\Bigr\}+\varepsilon$$
$$\leq \sup_{a,b\in[-t,s]:\,|a-b|\leq\delta}\bigl(-H^{a,b}_{s,t}(r-2\varepsilon)\bigr)\vee(-B\varepsilon/2)+\varepsilon.$$

We first let $\delta\searrow 0$, and by Lemma 4.4 the bound above becomes

$$-R_s(r)\leq\sup_{a\in[-t,s]}\bigl(-H^{a,a}_{s,t}(r-2\varepsilon)\bigr)\vee(-B\varepsilon/2)+\varepsilon.$$

Next we take $B\nearrow\infty$, and finally $\varepsilon\searrow 0$ with another application of Lemma 4.4. This establishes $\geq$ in (4.28). $\square$

A key analytic trick will be to look at the dual $J^*_{(t,t)-\bar{\mathbf{v}}(a)}(\xi)$ of the right tail rate as a function of a. This lemma will be helpful.

LEMMA 4.6. *For a fixed $\xi\in[0,\mu)$, the function*

$$G_\xi(a)=\begin{cases}-J^*_{(t,t)-\bar{\mathbf{v}}(a)}(\xi), & a\in[0,t],\\ \infty, & a<0 \text{ or } a>t,\end{cases}\tag{4.32}$$

is continuous on $[0,t]$, and convex and lower semi-continuous on $\mathbb{R}$. In particular, $G^{**}_\xi(a)=G_\xi(a)$ for $a\in\mathbb{R}$.

PROOF. To show convexity on $[0,t]$, let $\lambda\in(0,1)$ and $a=\lambda a_1+(1-\lambda)a_2$:

$$-J^*_{(t,t)-\bar{\mathbf{v}}(a)}(\xi)$$
$$=-\sup_{r\in\mathbb{R}}\{\xi r-J_{(t,t)-\bar{\mathbf{v}}(a)}(r)\}$$
$$=\inf_{r\in\mathbb{R}}\{J_{t-a,t}(r)-\xi r\}$$
$$\leq\inf_{r\in\mathbb{R}}\inf_{\substack{(r_1,r_2):\\ \lambda r_1+(1-\lambda)r_2=r}}\{\lambda\bigl(J_{t-a_1,t}(r_1)-\xi r_1\bigr)+(1-\lambda)\bigl(J_{t-a_2,t}(r_2)-\xi r_2\bigr)\}\tag{4.33}$$

$$
\begin{aligned}
&= \inf_{(r_1,r_2)\in\mathbb{R}^2}\{\lambda(J_{t-a_1,t}(r_1)-\xi r_1)+(1-\lambda)(J_{t-a_2,t}(r_2)-\xi r_2)\}\\
&= \lambda \inf_{r_1\in\mathbb{R}}\{J_{t-a_1,t}(r_1)-\xi r_1\}+(1-\lambda)\inf_{r_2\in\mathbb{R}}\{J_{t-a_2,t}(r_2)-\xi r_2\}\\
&= -\lambda J^*_{t-a_1,t}(\xi)-(1-\lambda)J^*_{t-a_2,t}(\xi).
\end{aligned}
$$

The inequality comes from the convexity of J in the variable $(t-a,t,r)$.

For finiteness on $[0,t]$ it is now enough to show that $G_\xi(a)$ is finite at the endpoints. Continuity then follows in the interior $(0,t)$. First take $a=t$. Then $J^*_{0,t}$ is the dual of a Cramér rate function, and for $\xi\ge 0$

$$G_\xi(t) = -J^*_{0,t}(\xi) = -t\log\mathbb{E}e^{\xi\log Y_{1,0}}, \tag{4.34}$$

which is finite for $\xi<\mu$.

Convexity of $J_{s,t}(r)$ and symmetry $J_{s,t}(r)=J_{t,s}(r)$ imply $J_{t,t}(r)\le J_{0,2t}(r)$.

From this

$$
\begin{aligned}
G_\xi(0) &= -J^*_{t,t}(\xi) = \inf_{r\in\mathbb{R}}\{J_{t,t}(r)-\xi r\}\\
&\le \inf_{r\in\mathbb{R}}\{J_{0,2t}(r)-\xi r\} = -J^*_{0,2t}(\xi)<\infty.
\end{aligned}
\tag{4.35}
$$

Continuity at $a=0$. To show that G_ξ is also continuous at the endpoints, we first obtain a lower bound. For any $r\in\mathbb{R}$,

$$J^*_{t-a,t}(\xi)\ge r\xi - J_{t-a,t}(r)$$

hence, by continuity of $J_{s,t}$ in the (s,t) argument,

$$\varliminf_{a\to 0} J^*_{t-a,t}(\xi)\ge r\xi - J_{t,t}(r). \tag{4.36}$$

Supremum over r gives $\varliminf_{a\to 0} J^*_{t-a,t}(\xi)\ge J^*_{t,t}(\xi)$.

For the upper bound, let $0<a<t$. Varadhan's theorem (Theorem 4.3.1 in [12]) applies in the present setting. This is justified in the proof of Corollary 2.5 below and another similar justification is given for (5.4) below. Consequently,

$$
\begin{aligned}
J^*_{t,t}(\xi) &= \lim_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z_{\lfloor nt\rfloor,\lfloor nt\rfloor}}\\
&\ge \lim_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z_{\lfloor n(t-a)\rfloor,\lfloor nt\rfloor}}\\
&\quad + \lim_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\sum_{i=\lfloor n(t-a)\rfloor+1}^{\lfloor nt\rfloor}\log Y_{i,\lfloor nt\rfloor}}\\
&= J^*_{t-a,t}(\xi)+a\log\mathbb{E}Y^\xi.
\end{aligned}
\tag{4.37}
$$

Taking $a\searrow 0$ yields continuity at $a=0$.

Continuity at $a=t$. The lower bound follows as in the previous case. For the upper bound we use a path counting argument. Let $e^{nF(s,t)}$ be an upper bound on the number of paths in $\Pi_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ such that $F(0+,t)=0$. Consider first the case where $0\le\xi<1$. Then

$$\begin{aligned} J^*_{t-a,t}(\xi) &= \lim_{n\to\infty} n^{-1}\log\mathbb{E}\Bigg(\sum_{x_\bullet\in\Pi_{(\lfloor n(t-a)\rfloor,\lfloor nt\rfloor)}}\prod_{i=1}^{\lfloor nt\rfloor+\lfloor n(t-a)\rfloor} Y_{x_i}\Bigg)^{\xi} \\ &\le \lim_{n\to\infty} n^{-1}\log\sum_{x_\bullet\in\Pi_{(\lfloor n(t-a)\rfloor,\lfloor nt\rfloor)}}\prod_{i=1}^{\lfloor nt\rfloor+\lfloor n(t-a)\rfloor}\mathbb{E}(Y)^{\xi} \\ &= F(t-a,t)+(2-a/t)J^*_{0,t}(\xi). \end{aligned} \tag{4.38}$$

For $1\le\xi<\mu$, Jensen's inequality yields

$$J^*_{t-a,t}(\xi)\le\xi F(t-a,t)+(2-a/t)J^*_{0,t}(\xi). \tag{4.39}$$

Let $a\nearrow t$ to get the continuity.

$G^{**}_\xi=G_\xi$ is a consequence of convexity and lower semicontinuity, by [28], Theorem 12.2. $\square$

PROOF OF PROPOSITION 4.3. The remainder of the proof is convex analysis. The goal is to derive the following formula for the right tail rate function $J_{s,t}$:

$$J_{s,t}(r)=\sup_{\xi\in[0,\mu)}\Big\{r\xi-\inf_{\theta\in(\xi,\mu)}\big(tM_\theta(\xi)-sM_{\mu-\theta}(-\xi)\big)\Big\}. \tag{4.40}$$

We begin by expressing the explicitly known dual $R^*_s(\xi)$ from (4.20) in terms of the unknown function $J_{(s,t)-\bar{\mathbf{v}}(a)}$. Equation (4.26) says that $H^a_{s,t}$ is the infimal convolution of κ_a and $J_{(s,t)-\bar{\mathbf{v}}(a)}$, in symbols $H^a_{s,t}=\kappa_a\Box J_{(s,t)-\bar{\mathbf{v}}(a)}$. By Theorem 16.4 in [28] addition is dual to infimal convolution. Starting with (4.28) we have

$$\begin{aligned} R^*_s(\xi) &= \sup_{-t\le a\le s}\sup_{r\in\mathbb{R}}\big\{r\xi-(\kappa_a\Box J_{(s,t)-\bar{\mathbf{v}}(a)})(r)\big\} \\ &= \sup_{-t\le a\le s}(\kappa_a\Box J_{(s,t)-\bar{\mathbf{v}}(a)})^*(\xi) \\ &= \sup_{-t\le a\le s}\big\{\kappa^*_a(\xi)+J^*_{(s,t)-\bar{\mathbf{v}}(a)}(\xi)\big\}. \end{aligned} \tag{4.41}$$

Combining this with (4.20) gives, for $0\le\xi<\theta$,

$$s\log\Gamma(\theta-\xi)-s\log\Gamma(\theta)=\sup_{-t\le a\le s}\big\{\kappa^*_a(\xi)+J^*_{(s,t)-\bar{\mathbf{v}}(a)}(\xi)\big\}. \tag{4.42}$$

Now regard $\xi\in[0,\mu)$ fixed, and let $\theta\in(\xi,\mu)$ vary. Introduce temporary definitions

$$u_a(\theta)=\begin{cases} -h_\xi(\theta)=M_{\mu-\theta}(-\xi)=\log\Gamma(\mu-\theta+\xi)-\log\Gamma(\mu-\theta), & -t\le a\le 0, \\ d_\xi(\theta)=M_\theta(\xi)=\log\Gamma(\theta-\xi)-\log\Gamma(\theta), & 0<a\le s. \end{cases} \tag{4.43}$$

Substitute (4.22) and (4.43) into equation (4.42) to get

$$s\log\frac{\Gamma(\theta-\xi)}{\Gamma(\theta)}-t\log\frac{\Gamma(\mu-\theta+\xi)}{\Gamma(\mu-\theta)}=\sup_{-t\le a\le s}\bigl\{au_a(\theta)+J^*_{(s,t)-\bar{\mathbf{v}}(a)}(\xi)\bigr\}. \tag{4.44}$$

The right-hand side begins to resemble a convex dual, and will allow us to solve for $J_{s,t}$. We can specialize to the case $s=t$ because $(t,t)-\bar{\mathbf{v}}(a)$ gives all the pairs (s,t) with $0\le s\le t$. When $s=t$, the $J_{s,t}=J_{t,s}$ symmetry allows us to write (4.44) as

$$t\bigl(d_\xi(\theta)+h_\xi(\theta)\bigr)=\sup_{0\le a\le t}\bigl\{a\bigl(h_\xi(\theta)\vee d_\xi(\theta)\bigr)+J^*_{t-a,t}(\xi)\bigr\},$$

and it splits into cases as follows:

$$t\bigl(d_\xi(\theta)+h_\xi(\theta)\bigr)=\begin{cases}\sup\limits_{0\le a\le t}\bigl\{ah_\xi(\theta)+J^*_{t-a,t}(\xi)\bigr\}, & \theta\in\bigl[(\mu+\xi)/2,\mu\bigr),\\ \sup\limits_{0\le a\le t}\bigl\{ad_\xi(\theta)+J^*_{t-a,t}(\xi)\bigr\}, & \theta\in\bigl(\xi,(\mu+\xi)/2\bigr].\end{cases}$$

We can discard one of the branches above. For if $\theta'=\mu+\xi-\theta$, then $d_\xi(\theta')=h_\xi(\theta)$, and we see that the two equations given by the two branches are in fact equivalent. So we restrict to the case $\theta\in[(\mu+\xi)/2,\mu)$ and continue with

$$t\bigl(d_\xi(\theta)+h_\xi(\theta)\bigr)=\sup_{0\le a\le t}\bigl\{ah_\xi(\theta)+J^*_{t-a,t}(\xi)\bigr\}. \tag{4.45}$$

The function h_ξ is strictly increasing, so we can change variables via $v=h_\xi(\theta)$ between the intervals $\theta\in[(\mu+\xi)/2,\mu)$ and $v\in[h_\xi((\mu+\xi)/2),\infty)$. Recall also $G_\xi(a)=-J^*_{t-a,t}(\xi)$ from Lemma 4.6. This turns (4.45) into

$$\begin{aligned}t\bigl((d_\xi\circ h_\xi^{-1})(v)+v\bigr)&=\sup_{0\le a\le t}\bigl\{av-G_\xi(a)\bigr\}\\ &=G^*_\xi(v),\qquad h_\xi\Bigl(\frac{\mu+\xi}{2}\Bigr)\le v<\infty.\end{aligned} \tag{4.46}$$

Utilizing $G_\xi=G^{**}_\xi$, we get the following expression for the rate function J:

$$J_{t-a,t}(r)=\sup_{\xi\in[0,\mu)}\bigl\{r\xi-J^*_{t-a,t}(\xi)\bigr\}=\sup_{\xi\in[0,\mu)}\bigl\{r\xi+G_\xi(a)\bigr\} \tag{4.47}$$

$$=\sup_{\xi\in[0,\mu)}\Bigl\{r\xi+\sup_{v\in\mathbb{R}}\bigl[av-G^*_\xi(v)\bigr]\Bigr\}$$

$$=\sup_{\xi\in[0,\mu)}\Bigl\{r\xi+\sup_{v\in[h_\xi((\mu+\xi)/2),\infty)}\bigl[av-G^*_\xi(v)\bigr]\Bigr\}$$

$$=\sup_{\xi\in[0,\mu)}\Bigl\{r\xi+\sup_{v\in[h_\xi((\mu+\xi)/2),\infty)}\bigl[(a-t)v-td_\xi\bigl(h_\xi^{-1}(v)\bigr)\bigr]\Bigr\}. \tag{4.48}$$

In the next to last equality above, we restricted the supremum over v to the interval $v\in[h_\xi((\mu+\xi)/2),\infty)$. This is justified because G^*_ξ is convex, $a\ge 0$ and from

(4.46) we can compute the right derivative $(G^*_\xi)'(h_\xi(\frac{\mu+\xi}{2})+) = 0$. The restriction of the supremum then allows us to replace $G^*_\xi(v)$ with (4.46).

The proof is complete. In the case $0 < s \le t$, take $a = t - s$ on line (4.47). Line (4.48) is the desired representation for $J_{s,t}$. It turns into (4.40) by the v to θ change of variable. The case $s > t$ follows from the symmetry $J_{s,t}(r) = J_{t,s}(r)$. $\square$

The next lemma makes explicit the formula(s) for $J^*_{s,t}$ that were implicit in the proof of Proposition 4.3.

LEMMA 4.7. *Let $s, t \ge 0$ and $\xi \in [0, \mu)$. Then*

$$J^*_{s,t}(\xi) = \inf_{\rho \in (\xi,\mu)} \{t M_\rho(\xi) - s M_{\mu-\rho}(-\xi)\} \tag{4.49}$$

$$= \inf_{\theta \in (\xi,\mu)} \{s M_\theta(\xi) - t M_{\mu-\theta}(-\xi)\}. \tag{4.50}$$

PROOF. (4.50) comes from (4.49) by the change of variable $\rho = \mu + \xi - \theta$. Comparison of the two shows that we can assume $s \le t$. To prove (4.49) for $s \le t$, start from Lemma 4.6:

$$J^*_{s,t}(\xi) = -G_\xi(t-s) = -G^{**}_\xi(t-s) = -\sup_{v \in \mathbb{R}} \{(t-s)v - G^*_\xi(v)\}.$$

Restrict the supremum as in (4.47) and (4.48), substitute in (4.46) and change variables from v to $\theta = h_\xi^{-1}(v)$. $\square$

PROOF OF COROLLARY 2.5. If $\xi \ge \mu$,

$$\xi \log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor} \ge \sum_j \xi \log Y_{x_j}$$

for any particular path $x_\centerdot \in \Pi_{\lfloor ns \rfloor, \lfloor nt \rfloor}$, and then $\Lambda_{s,t}(\xi) = \infty$ comes from $M_\mu(\xi) = \infty$ from (2.2).

Let $\xi < \mu$. Pick $\gamma > 1$ such that $\gamma\xi < \mu$. Then the bound

$$\sup_n n^{-1} \log \mathbb{E} e^{\gamma \xi \log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor}} < \infty$$

follows from path counting, as in (4.38) and (4.39). This bound is sufficient for Varadhan's theorem (Theorem 4.3.1 in [12]) which gives

$$\begin{aligned} \lim_{n\to\infty} \Lambda_{s,t}(\xi) = n^{-1} \log \mathbb{E} e^{\xi \log Z_{\lfloor ns \rfloor, \lfloor nt \rfloor}} &= I^*_{s,t}(\xi) = \sup_{r \in \mathbb{R}} \{r\xi - I_{s,t}(r)\} \\ &= \sup_{r \ge p(s,t)} \{r\xi - I_{s,t}(r)\} = \sup_{r \ge p(s,t)} \{r\xi - J_{s,t}(r)\}. \end{aligned}$$

We discarded $\{I_{s,t} = \infty\} = \{r < p(s,t)\}$ from the supremum. Since $I_{s,t}$ increases for $r \ge p(s,t)$, the case $\xi \le 0$ of (2.11) follows. For $\xi \ge 0$ the values $J_{s,t}(r) = 0$

for $r < p(s,t)$ can be put back in because they do not alter the supremum. Consequently $\Lambda_{s,t}(\xi) = J^*_{s,t}(\xi)$ for $\xi \ge 0$. Lemma 4.7 completes the proof of this corollary. $\square$

There is nothing new in the proof of Corollary 2.8, so we omit it.

5. Proofs for the stationary log-gamma model. In this section we prove the results of Section 2.2.

PROOF OF THEOREM 2.14. Coarse-graining arguments and simple error bounds readily give the following limit:

$$\begin{aligned} p^{(\theta),\mathrm{hor}}(s,t) &= \lim_{n\to\infty} n^{-1} \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \\ &= \lim_{n\to\infty} \max_{1\le k\le \lfloor ns\rfloor} \left(n^{-1} \sum_{i=1}^{k} \log U_{i,0} + n^{-1} \log Z^{\square}_{(k,1),(\lfloor ns\rfloor,\lfloor nt\rfloor)} \right) \\ &= \sup_{0\le a\le s} \{-a\Psi_0(\theta) + p(s-a,t)\} \\ &= \sup_{0\le a\le s} \inf_{0<\rho<\mu} \{-a\Psi_0(\theta) + (a-s)\Psi_0(\rho) - t\Psi_0(\mu-\rho)\}. \end{aligned}$$

In the last step we substituted in (2.5). Formula (2.30) follows from this by some calculus.

From the definition (2.22) of $Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$, follow inequalities analogous to (4.17), and then with arguments like those in the proof of Lemma 4.5, we derive a right tail LDP

$$\begin{aligned} &\lim_{n\to\infty} n^{-1} \log \mathbb{P}\{Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \ge nr\} \\ &\qquad = -J_{\theta,\mathrm{hor}}(r) = -\inf_{a\in[0,s]} (R_a \square J_{s-a,t})(r), \end{aligned} \tag{5.1}$$

where R_a is the rate function from (4.18). For $\xi \ge 0$ the l.m.g.f. in (2.24) satisfies $\Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) = J^*_{\theta,\mathrm{hor}}(\xi)$. This would be a consequence of Varadhan's theorem if we had a full LDP, but now we have to justify this separately, and we do so in Lemma 5.1 below. 1 Proceeding as in (4.41) and using (4.50),

$$\begin{aligned} \Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) &= \sup_{a\in[0,s]} \bigl(R^*_a(\xi) + J^*_{s-a,t}(\xi)\bigr) \\ &= \sup_{a\in[0,s]} \inf_{\rho\in(\xi,\mu)} \{aM_\theta(\xi) + (s-a)M_\rho(\xi) - tM_{\mu-\rho}(-\xi)\}. \end{aligned}$$

Formula (2.31) follows from some calculus. The sup and inf can be interchanged by a minimax theorem (see, e.g., [19]), and this makes the calculus easier. $\square$

LEMMA 5.1. *Let $Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$ the partition function given by* (2.22), *and let $J_{\theta,\mathrm{hor}}(r)$ as given by* (5.1). *Then for $0\le\xi<\theta$,*

$$\lim_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}=\sup_{r\in\mathbb{R}}\{r\xi-J_{\theta,\mathrm{hor}}(r)\}=J^*_{\theta,\mathrm{hor}}(\xi).$$

PROOF. Let $0<\xi<\theta$. Set

$$\underline{\gamma}=\varliminf_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\quad\text{and}\quad\overline{\gamma}=\varlimsup_{n\to\infty} n^{-1}\log\mathbb{E}e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}.$$

First we have an exponential Chebyshev argument for a lower bound:

$$n^{-1}\log\mathbb{P}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\ge nr\}\le-\xi r+n^{-1}\log\mathbb{E}e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}.$$

Letting $n\to\infty$ along a suitable subsequence gives $\underline{\gamma}\ge\xi r-J_{\theta,\mathrm{hor}}(r)$ for all $r\in\mathbb{R}$. Thus $\underline{\gamma}\ge J^*_{\theta,\mathrm{hor}}(\xi)$ holds.

For the upper bound we claim that

$$\lim_{r\to\infty}\varlimsup_{n\to\infty} n^{-1}\log\mathbb{E}\bigl(e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\mathbf{1}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\ge nr\}\bigr)=-\infty.\tag{5.2}$$

Assume for a moment that (5.2) holds. To establish the upper bound let $\delta>0$ and partition $\mathbb{R}$ with $r_i=i\delta$, $i\in\mathbb{Z}$:

$$\begin{aligned}
&n^{-1}\log\mathbb{E}\bigl(e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\bigr)\\
&\quad\le n^{-1}\log\Biggl[\sum_{i=-m}^{m}e^{n\xi r_{i+1}}\mathbb{P}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\ge nr_i\}\\
&\qquad\qquad+e^{n\xi r_{-m}}+\mathbb{E}\bigl(e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\mathbf{1}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\ge nr_m\}\bigr)\Biggr].
\end{aligned}\tag{5.3}$$

By (5.2), for each $M>0$ there exists $m=m(M)$ so that

$$n^{-1}\log\mathbb{E}\bigl(e^{\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\mathbf{1}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}\ge nr_m\}\bigr)<-M.$$

A limit along a suitable subsequence in (5.3) yields

$$\begin{aligned}
\overline{\gamma}&\le\max_{-m\le i\le m}\{\xi r_{i+1}-J_{\theta,\mathrm{hor}}(r_i)\}\vee\xi r_{-m}\vee(-M)\\
&\le\Bigl(\sup_{r\in\mathbb{R}}\{\xi r-J_{\theta,\mathrm{hor}}(r)\}+\xi\delta\Bigr)\vee\xi r_{-m}\vee(-M).
\end{aligned}$$

The proof of the lemma follows by letting $\delta\to0$, $m\to\infty$ and $M\to\infty$.

Now to show (5.2). Note that there exists $\alpha>1$ such that $\alpha\xi<\theta$,

$$\sup_n\bigl(\mathbb{E}e^{\alpha\xi\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\bigr)^{1/n}<\infty.\tag{5.4}$$

To see this, distinguish cases where $\alpha\xi < 1$ or otherwise. Let N denote the number of paths, and recall that $N \le e^{cn}$ for some $c > 0$: For $\alpha\xi < 1$,

$$\big(\mathbb{E}e^{\alpha\xi \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\big)^{1/n} = \left(\mathbb{E}\left[\left(\sum_{x\in\Pi_{(\lfloor ns\rfloor,\lfloor nt\rfloor)}} \prod_{i=1}^{\lfloor ns\rfloor+\lfloor nt\rfloor} Y_{x_j}\right)^{\alpha\xi}\right]\right)^{1/n}$$

$$\le \left(N \prod_{i=1}^{\lfloor nt\rfloor+\lfloor ns\rfloor} \mathbb{E}Y^{\alpha\xi}\right)^{1/n} \le e^c M_\theta(\alpha\xi)^{t+s}.$$

For $\alpha\xi \ge 1$, Jensen's inequality gives

$$\big(\mathbb{E}e^{\alpha\xi \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\big)^{1/n} = \left(\mathbb{E}\left[\left(\sum_{x\in\Pi_{(\lfloor ns\rfloor,\lfloor nt\rfloor)}} \prod_{i=1}^{\lfloor ns\rfloor+\lfloor nt\rfloor} Y_{x_j}\right)^{\alpha\xi}\right]\right)^{1/n}$$

$$\le \left(N^{\alpha\xi} \prod_{i=1}^{\lfloor nt\rfloor+\lfloor ns\rfloor} \mathbb{E}Y^{\alpha\xi}\right)^{1/n} \varepsilon e^{c\alpha\xi} M_\theta(\alpha\xi)^{t+s}.$$

To show (5.2), use Hölder's inequality,

$$\begin{aligned} &n^{-1}\log \mathbb{E}\big(e^{\xi \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}} \mathbf{1}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \ge nr\}\big)\\ &\quad\le \alpha^{-1}\log\sup_n \big(\mathbb{E}e^{\alpha\xi \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}}\big)^{1/n}\\ &\qquad + (\alpha-1)\alpha^{-1}n^{-1}\log \mathbb{P}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \ge nr\}. \end{aligned}$$

Taking a limit $n \to \infty$, we conclude

$$\varlimsup_{n\to\infty} n^{-1}\log \mathbb{E}\big(e^{\xi \log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}} \mathbf{1}\{\log Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor} \ge nr\}\big) \le C_1 - C_2 J_{\theta,\mathrm{hor}}(r) \tag{5.5}$$

for positive constants C_1, C_2. Letting $r \to \infty$ finishes the proof because

$$\lim_{r\to\infty} J_{\theta,\mathrm{hor}}(r) = \infty. \qquad \square$$

PROOF OF THEOREM 2.11. We can assume $0 < \xi < \theta \wedge (\mu - \theta)$ because otherwise the boundary variables alone force the l.m.g.f. to blow up.

Let us record the counterpart of (2.31) for $Z^{(\theta),\mathrm{hor}}_{\lfloor ns\rfloor,\lfloor nt\rfloor}$. Condition (2.28) becomes

$$t\big(\Psi_0(\mu-\theta) - \Psi_0(\mu-\theta-\xi)\big) \ge s\big(\Psi_0(\theta+\xi) - \Psi_0(\theta)\big). \tag{5.6}$$

The conclusion becomes that the limit in (2.25) exists and is given by

$$\Lambda^{\mathrm{ver}}_{\theta,(s,t)}(\xi) = \begin{cases} tM_{\mu-\theta}(\xi) - sM_\theta(-\xi), & \text{if (5.6) holds,}\\ \Lambda_{t,s}(\xi) = \Lambda_{s,t}(\xi), & \text{if (5.6) fails.}\end{cases} \tag{5.7}$$

The logarithmic limits lead to the formula

$$\Lambda_{\theta,(s,t)}(\xi) = \Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) \vee \Lambda^{\mathrm{ver}}_{\theta,(s,t)}(\xi), \tag{5.8}$$

and we need to justify that this is the same as the maximum in (2.20). This comes from several observations:

(i) $\Lambda_{s,t}(\xi) = J^*_{s,t}(\xi)$ is always bounded above by the first branches of both (2.31) and (5.7). This is evident from equations (4.49) and (4.50).

(ii) Conditions (2.28) and (5.6) together define three ranges for (s, t):
(a) (2.28) and (5.6) both hold if and only if $\alpha_1 t \le s \le \alpha_2 t$;
(b) (2.28) holds and (5.6) fails if and only if $s > \alpha_2 t$;
(c) (2.28) fails and (5.6) holds if and only if $s < \alpha_1 t$.
The constants $0 < \alpha_1 < \alpha_2$ can be read off (2.28) and (5.6), and the strict inequalities are justified by the strict concavity of Ψ_0.

(iii) In the maximum in (2.20), we have

$$sM_\theta(\xi) - tM_{\mu-\theta}(-\xi) \ge tM_{\mu-\theta}(\xi) - sM_\theta(-\xi) \tag{5.9}$$

if and only if $s \ge \alpha_3 t$ for a constant $\alpha_3 > 0$ that can be read off from above. Strict concavity of Ψ_0 implies that $0 < \alpha_1 < \alpha_3 < \alpha_2$.

Now we argue that

$$\Lambda_{\theta,(s,t)}(\xi) = \max\{sM_\theta(\xi) - tM_{\mu-\theta}(-\xi), tM_{\mu-\theta}(\xi) - sM_\theta(-\xi)\}. \tag{5.10}$$

This is clear in case (a) as this maximum is exactly $\Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi) \vee \Lambda^{\mathrm{ver}}_{\theta,(s,t)}(\xi)$. In case (b), $\Lambda^{\mathrm{hor}}_{\theta,(s,t)}(\xi)$ equals the left-hand side of (5.9) which dominates both the right-hand side of (5.9) and $\Lambda_{s,t}(\xi)$. Consequently in case (b) also (5.8) is the same as (5.10). Case (c) is symmetric to (b). This completes the proof of (5.10).

With one additional observation we can verify Remark 2.15. Namely, $\Lambda_{s,t}(\xi)$ is in fact *strictly* bounded above by the first branch of either (2.31) or (5.7). The claim is easily verifiable when either of conditions (b) or (c) are in effect. To see the strict domination when (a) holds, note that the unique minimizers in formulas (4.49) and (4.50) are linked by $\rho = \mu + \xi - \theta$. But if these formulas matched both first branches in (2.31) and (5.7), the connection would have to be $\rho = \mu - \theta$. This together with (5.10) implies that $\Lambda_{s,t}(\xi) < \Lambda_{\theta,(s,t)}(\xi)$ for all $\theta \in (0, \mu)$. □

Acknowledgments. We would like to thank an anonymous referee for a thorough reading of the original manuscript and suggestions that improved the presentation of this article.

REFERENCES

[1] AMIR, G., CORWIN, I. and QUASTEL, J. (2011). Probability distribution of the free energy of the continuum directed random polymer in $1+1$ dimensions. *Comm. Pure Appl. Math.* **64** 466–537. MR2796514

[2] BEN-ARI, I. (2009). Large deviations for partition functions of directed polymers in an IID field. *Ann. Inst. Henri Poincaré Probab. Stat.* **45** 770–792. MR2548503

[3] BOLTHAUSEN, E. (1989). A note on the diffusion of directed polymers in a random environment. *Comm. Math. Phys.* **123** 529–534. MR1006293

[4] CARMONA, P. and HU, Y. (2002). On the partition function of a directed polymer in a Gaussian random environment. *Probab. Theory Related Fields* **124** 431–457. MR1939654

[5] COMETS, F., SHIGA, T. and YOSHIDA, N. (2003). Directed polymers in a random environment: Path localization and strong disorder. *Bernoulli* **9** 705–723. MR1996276

[6] COMETS, F., SHIGA, T. and YOSHIDA, N. (2004). Probabilistic analysis of directed polymers in a random environment: A review. In *Stochastic Analysis on Large Scale Interacting Systems. Advanced Studies in Pure Mathematics* **39** 115–142. Math. Soc. Japan, Tokyo. MR2073332

[7] COMETS, F. and VARGAS, V. (2006). Majorizing multiplicative cascades for directed polymers in random media. *ALEA Lat. Am. J. Probab. Math. Stat.* **2** 267–277. MR2249671

[8] COMETS, F. and YOSHIDA, N. (2006). Directed polymers in random environment are diffusive at weak disorder. *Ann. Probab.* **34** 1746–1770. MR2271480

[9] COMETS, F. and YOSHIDA, N. (2011). Branching random walks in space–time random environment: Survival probability, global and local growth rates. *J. Theoret. Probab.* **24** 657–687. MR2822477

[10] CORWIN, I. (2011). The Kardar–Parisi–Zhang equation and universality class. Available at arXiv:1106.1596.

[11] CORWIN, I., O'CONNELL, N., SEPPÄLÄINEN, T. and ZYGOURAS, N. (2011). Tropical combinatorics and Whittaker functions. Available at arXiv:1110.3489.

[12] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036

[13] DEN HOLLANDER, F. (2009). *Random Polymers. Lecture Notes in Math.* **1974**. Springer, Berlin. Lectures from the 37th Probability Summer School held in Saint-Flour, 2007. MR2504175

[14] DEUSCHEL, J.-D. and ZEITOUNI, O. (1999). On increasing subsequences of I.I.D. samples. *Combin. Probab. Comput.* **8** 247–263. MR1702546

[15] GEORGIOU, N. (2011). Positive and zero temperature polymer models. Ph.D. thesis, Univ. Wisconsin–Madison. Available at arXiv:1210.0600.

[16] HUSE, D. A. and HENLEY, C. L. (1985). Pinning and roughening of domain wall in Ising systems due to random impurities. *Phys. Rev. Lett.* **54** 2708–2711.

[17] IMBRIE, J. Z. and SPENCER, T. (1988). Diffusion of directed polymers in a random environment. *J. Stat. Phys.* **52** 609–626. MR0968950

[18] JOHANSSON, K. (2000). Shape fluctuations and random matrices. *Comm. Math. Phys.* **209** 437–476. MR1737991

[19] KASSAY, G. (1994). A simple proof for König's minimax theorem. *Acta Math. Hungar.* **63** 371–374. MR1261480

[20] KIM, J. H. (1996). On increasing subsequences of random permutations. *J. Combin. Theory Ser. A* **76** 148–155. MR1405997

[21] LACOIN, H. (2010). New bounds for the free energy of directed polymers in dimension $1+1$ and $1+2$. *Comm. Math. Phys.* **294** 471–503. MR2579463

[22] LIU, Q. and WATBLED, F. (2009). Exponential inequalities for martingales and asymptotic properties of the free energy of directed polymers in a random environment. *Stochastic Process. Appl.* **119** 3101–3132. MR2568267

[23] LOGAN, B. F. and SHEPP, L. A. (1977). A variational problem for random Young tableaux. *Adv. Math.* **26** 206–222. MR1417317

[24] MORENO, G. (2010). Convergence of the law of the environment seen by the particle for directed polymers in random media in the L^2 region. *J. Theoret. Probab.* **23** 466–477. MR2644870

[25] MORIARTY, J. and O'CONNELL, N. (2007). On the free energy of a directed polymer in a Brownian environment. *Markov Process. Related Fields* **13** 251–266. MR2343849

[26] O'CONNELL, N. (2012). Directed polymers and the quantum Toda lattice. *Ann. Probab.* **40** 437–458.

[27] O'CONNELL, N. and YOR, M. (2001). Brownian analogues of Burke's theorem. *Stochastic Process. Appl.* **96** 285–304. MR1865759

[28] ROCKAFELLAR, R. T. (1970). *Convex Analysis. Princeton Mathematical Series* **28**. Princeton Univ. Press, Princeton, NJ. MR0274683

[29] SEPPÄLÄINEN, T. (1998). Coupling the totally asymmetric simple exclusion process with a moving interface. *Markov Process. Related Fields* **4** 593–628. I Brazilian School in Probability (Rio de Janeiro, 1997). MR1677061

[30] SEPPÄLÄINEN, T. (1998). Large deviations for increasing sequences on the plane. *Probab. Theory Related Fields* **112** 221–244. MR1653841

[31] SEPPÄLÄINEN, T. (2012). Scaling for a one-dimensional directed polymer with boundary conditions. *Ann. Probab.* **40** 19–73. MR2917766

[32] SEPPÄLÄINEN, T. and VALKÓ, B. (2010). Bounds for scaling exponents for a $1+1$ dimensional directed polymer in a Brownian environment. *ALEA Lat. Am. J. Probab. Math. Stat.* **7** 451–476. MR2741194

[33] WATBLED, F. (2012). Concentration inequalities for disordered models. *ALEA Lat. Am. J. Probab. Math. Stat.* **9** 129–140. MR2904479

MATHEMATICS DEPARTMENT
UNIVERSITY OF UTAH
JWB
155 S 1400 E
SALT LAKE CITY, UTAH 84112-0090
USA
E-MAIL: georgiou@math.utah.edu
URL: http://www.math.utah.edu/~georgiou

MATHEMATICS DEPARTMENT
UNIVERSITY OF WISCONSIN–MADISON
VAN VLECK HALL
480 LINCOLN DR.
MADISON, WISCONSIN 53706-1388
USA
E-MAIL: seppalai@math.wisc.edu
URL: http://www.math.wisc.edu/~seppalai