The Annals of Applied Probability
1993, Vol. 3, No. 4, 1067–1099

A VARIATIONAL CHARACTERIZATION OF THE SPEED OF A ONE-DIMENSIONAL SELF-REPELLENT RANDOM WALK

BY ANDREAS GREVEN AND FRANK DEN HOLLANDER

Universität Göttingen and Universiteit Utrecht

Let Q_n^α be the probability measure for an n-step random walk $(0, S_1, \ldots, S_n)$ on $\mathbb{Z}$ obtained by weighting simple random walk with a factor $1 - \alpha$ for every self-intersection. This is a model for a one-dimensional polymer. We prove that for every $\alpha \in (0,1)$ there exists $\theta^*(\alpha) \in (0,1)$ such that

$$\lim_{n\to\infty} Q_n^\alpha\left(\frac{|S_n|}{n} \in [\theta^*(\alpha) - \varepsilon, \theta^*(\alpha) + \varepsilon]\right) = 1 \quad \text{for every } \varepsilon > 0.$$

We give a characterization of $\theta^*(\alpha)$ in terms of the largest eigenvalue of a one-parameter family of $\mathbb{N} \times \mathbb{N}$ matrices. This allows us to prove that $\theta^*(\alpha)$ is an analytic function of the strength α of the self-repellence. In addition to the speed we prove a limit law for the local times of the random walk. The techniques used enable us to treat more general forms of self-repellence involving multiple intersections. We formulate a partial differential inequality that is equivalent to $\alpha \to \theta^*(\alpha)$ being (strictly) increasing. The verification of this inequality remains open.

0. Introduction.

0.1. *Motivation.* The purpose of this paper is to apply combinatorial and variational techniques developed in Greven and den Hollander (1992) and Baillon, Clément, Greven and den Hollander (1991) to answer some questions about a one-dimensional polymer model. The polymer is modelled probabilistically as a self-repellent random walk which pays a penalty $1 - \alpha$ for every self-intersection. The basic question is how fast this walk moves, which corresponds to the spread of the polymer as a function of its length. Bolthausen (1990) proved that if α is sufficiently small, then the displacement of the walk per unit of time is bounded between strictly positive finite constants. He raised the question: "Does the walk have an asymptotically deterministic speed? If so, then can this speed be computed and is it an analytic and strictly increasing function of the strength α of the self-repellence?" We shall answer his question in the affirmative, with the exception of the increasing property. Our results are consequences of a large deviation analysis for a Markov chain underlying the local times of the random walk. This leads to a representation via a largest eigenvalue problem for positive self-adjoint compact operators on $l^2(\mathbb{N})$.

Received August 1991; revised March 1993.
AMS 1991 *subject classifications.* Primary 60K35; secondary 58E30, 60F10, 60J15.
Key words and phrases. Polymer model, self-repellent random walk, large deviations, variational problems.

For a related model based on Brownian motion, see Westwater (1984) and Kusuoka (1985).

0.2. *Model.* Let $(S_i)_{i=0}^{\infty}$ be a simple random walk on $\mathbb{Z}$ starting at $S_0 = 0$ (i.e., the increments are i.i.d. $+1$ or -1 with probability $\frac{1}{2}$). Let P_n be the law of $(S_i)_{i=0}^{n}$ and let E_n be the corresponding expectation. Define a new law Q_n^{α} on n-step paths by setting

$$\frac{dQ_n^{\alpha}}{dP_n}\big((S_i)_{i=0}^{n}\big) = \frac{1}{Z_n^{\alpha}} \prod_{\substack{i,j=0\\ i\neq j}}^{n} \big(1 - \alpha 1\{S_i = S_j\}\big), \tag{0.1}$$

where

$$Z_n^{\alpha} = E_n\left(\prod_{\substack{i,j=0\\ i\neq j}}^{n} \big(1 - \alpha 1\{S_i = S_j\}\big)\right) \tag{0.2}$$

is the normalization factor and $\alpha \in (0,1)$. Q_n^{α} is the n-polymer measure with strength of repellence α.

0.3. *Main result for polymers.*

THEOREM 0. *There exists a function* $\theta^*\colon (0,1) \to (0,1)$ *such that*

$$\lim_{n\to\infty} Q_n^{\alpha}\left(\frac{|S_n|}{n} \in \big[\theta^*(\alpha) - \varepsilon, \theta^*(\alpha) + \varepsilon\big]\right) = 1 \quad \textit{for every } \varepsilon > 0, \tag{0.3}$$

$$\begin{aligned} &\alpha \to \theta^*(\alpha) \textit{ is analytic},\\ &\lim_{\alpha \downarrow 0} \theta^*(\alpha) = 0 \textit{ and } \lim_{\alpha \uparrow 1} \theta^*(\alpha) = 1. \end{aligned} \tag{0.4}$$

Later on we shall give a characterization of $\theta^*(\alpha)$ in terms of an eigenvalue problem [see (1.14)].

REMARK. It is natural to suspect that $\alpha \to \theta^*(\alpha)$ is (strictly) increasing. The characterization of $\theta^*(\alpha)$ will enable us to show that this property is equivalent to a second order partial differential inequality for the largest eigenvalue of a two-parameter family of $\mathbb{N} \times \mathbb{N}$ matrices. This fact indicates, in particular, that the increasing property is much more subtle than one might expect. We refer to Corollary $*$ and Section 3.6.

The techniques used actually allow us to prove Theorem 0 in the more general context where we define the polymer measure Q_n^{α} via the functional

$$\prod_{k\geq 2} \prod_{i_1,\ldots,i_k=0}^{n} \big(1 - \alpha_k 1\{S_{i_1} = \cdots = S_{i_k}\}\big), \tag{0.5}$$

where α is replaced by a sequence

$$\alpha = (\alpha_k)_{k=2}^{\infty}, \qquad \alpha_k \in [0,1) \text{ for } k \geq 2, \ \sum_{k \geq 2} \alpha_k > 0. \tag{0.6}$$

In this model multiple intersections get a penalty depending on their multiplicity.

0.4. *Local times.* It is more convenient for the analysis of the model to view the functional in (0.5) in a different light. Define the local time at site x of the n-step path by

$$l_n(x) = \sum_{i=0}^{n} 1\{S_i = x\}. \tag{0.7}$$

Then (0.5) may be rewritten as

$$\exp\left[-\sum_{x \in \mathbb{Z}} g(l_n(x))\right] \tag{0.8}$$

with

$$g(l) = \sum_{k \geq 2} \beta_k l^k, \qquad \beta_k = -\log(1-\alpha_k). \tag{0.9}$$

Thus we see that the weight factor is an exponential functional of the local times given by a nonlinear function g. The polymer measure in (0.1) is the special case where $g(l) = \beta l^2$. [Note that the terms in (0.1) with $i = j$ give a factor $(1-\alpha)^{n+1}$ which may be absorbed in Z_n^{α}.]

0.5. *Outline.* Our paper is organized as follows. In Section 1 we formulate three theorems. Theorem 1 gives a representation of the form

$$\lim_{n \to \infty} \frac{1}{n} \log Z_n^{\alpha} = \sup_{\theta \in (0,1]} \sup_{\nu \in M_\theta} [-\theta \phi^{\alpha}(\nu)], \tag{0.10}$$

where $\phi^{\alpha}(\nu)$ is some nonlinear functional of ν and M_θ is some subset of the probability measures on $\mathbb{N}^2$ given by a θ-dependent linear constraint. Here θ plays the role of the speed of the path and ν the role of the empirical distribution of the numbers of right jumps and left jumps in the path (counted along the stretch visited), both in the limit as $n \to \infty$. The value $-\theta\phi^{\alpha}(\nu)$ is the exponential rate at which paths corresponding to θ, ν contribute to Z_n^{α}. Theorem 2 solves the variational problems in (0.10) in terms of an eigenvalue problem for a one-parameter family of $\mathbb{N} \times \mathbb{N}$ positive symmetric matrices, which induce positive self-adjoint compact operators on $l^2(\mathbb{N})$. Theorem 3 formulates some consequences of Theorems 1 and 2 for the path properties under Q_n^{α} as $n \to \infty$. These path properties are directly related to the maximizers in (0.10).

The three theorems are formulated and proved in a somewhat more general context, and immediately imply Theorem 0. In Sections 2–4 we give the proofs of Theorems 1–3, respectively.

0.6. *Related models.* Squares of local times appear in various related models. A nice example is: Let $(\omega_x)_{x\in\mathbb{Z}}$ be i.i.d. with Gaussian density $p_\sigma(\omega)$ with mean zero and variance σ^2. Think of ω_x as the rotation speed of a magnetic spin at site x. Let

$$\Omega_n = \exp\left[i\sum_{j=0}^{n}\omega_{S_j}\right], \qquad i^2 = -1, \tag{0.11}$$

denote the total rotation angle of a spin that follows an n-step simple random walk with jumps at unit time intervals. If E_p denotes expectation over $(\omega_x)_{x\in\mathbb{Z}}$, then

$$\begin{aligned} E_n E_p(\Omega_0\Omega_n) &= E_n\left(\prod_x \int p_\sigma(\omega)\exp(i\omega l_n(x))\,d\omega\right) \\ &= E_n\left(\prod_x \exp\left(-\tfrac{1}{2}\sigma^2 l_n^2(x)\right)\right), \end{aligned} \tag{0.12}$$

which is the same as Z_n^α in (0.2) with $-\log(1-\alpha) = \frac{1}{2}\sigma^2$. This is a model for spin depolarization in nuclear magnetic resonance experiments [Czech and Kehr (1984), (1986) and Mazo and van den Broeck (1986)].

1. Main results.

1.1. *Representation as variational problems.* In this section we give a variational representation for the exponential growth rate of functionals of the type (0.8) under the law of simple random walk. Let $g\colon \mathbb{N}\cup\{0\}\to[0,\infty)$ be given with the properties

(1.1)(i) $\quad g(0) = 0,$

(1.1)(ii) $\quad g(i+j) > g(i) + g(j)\quad$ for all $i,j\in\mathbb{N}$,

(1.1)(iii) $\quad \lim_{i\to\infty} g(i)/i = \infty.$

Define the matrix A by

$$A(i,j) = e^{-g(i+j-1)}P(i,j), \qquad i,j\in\mathbb{N}, \tag{1.2}$$

with Markov transition matrix P given by

$$P(i,j) = \binom{i+j-2}{i-1}\left(\tfrac{1}{2}\right)^{i+j-1}, \qquad i,j\in\mathbb{N}. \tag{1.3}$$

Let $\mathscr{P}(\mathbb{N}^2)$ denote the set of probability measures ν on $\mathbb{N}^2$ and $\hat{\nu}$ the marginal $\hat{\nu}(i) = \sum_{j\in\mathbb{N}}\nu(i,j)$.

THEOREM 1. *For every g satisfying* (1.1),

$$\lim_{n\to\infty}\frac{1}{n}\log E_n\left(\exp\left[-\sum_{x\in\mathbb{Z}} g(l_n(x))\right]\right) = -z. \tag{1.4}$$

The quantity z is given by the variational problems

$$z = \inf_{\theta\in(0,1]} \theta K(\theta), \tag{1.5}$$

$$K(\theta) = \inf_{\nu\in M_\theta} \phi(\nu) \tag{1.6}$$

with

$$\phi(\nu) = \sum_{i,j\in\mathbb{N}} \nu(i,j)\log\left(\frac{\nu(i,j)}{\hat{\nu}(i)A(i,j)}\right), \tag{1.7}$$

$$M_\theta = \Big\{\nu\in\mathscr{P}(\mathbb{N}^2)\colon \sum_{j\in\mathbb{N}}\nu(i,j) = \sum_{j\in\mathbb{N}}\nu(j,i) \text{ for all } i\in\mathbb{N}, \quad \sum_{i,j\in\mathbb{N}}(i+j-1)\nu(i,j) = \theta^{-1}\Big\}. \tag{1.8}$$

The kernel P in (1.9) appears because the local times of a simple random walk can be viewed as a functional of the Markov chain generated by P. For more on this, see Section 3.

1.2. *Solution of variational problems.* We now give the solution of (1.5) and (1.6). The minimum and the minimizers will be given in terms of the largest eigenvalue and corresponding eigenvector of the matrix A_r defined by

$$A_r(i,j) = e^{r(i+j-1)}A(i,j), \qquad i,j\in\mathbb{N},\ r\in\mathbb{R}. \tag{1.9}$$

For every r, $A_r(i,j) = A_r(j,i) > 0$ for all i,j and $\sum_{i,j}A_r^2(i,j) < \infty$ [use (1.1)(iii)]. Hence the operator $A_r\colon l^2(\mathbb{N})\to l^2(\mathbb{N})$ is positive self-adjoint and compact, and therefore there exists a unique pair $(\lambda(r),\tau_r)$ such that

$$\begin{aligned}&\lambda(r)\tau_r = A_r\tau_r,\\ &\lambda(r)\in(0,\infty),\ \tau_r\in l^2(\mathbb{N}),\ \|\tau_r\|_2 = 1,\ \tau_r>0,\\ &r\to\lambda(r) \text{ and } r\to\tau_r \text{ are analytic}\end{aligned} \tag{1.10}$$

[compare Baillon, Clément, Greven and den Hollander (1991), Sections 3.2 and 3.3]. Moreover, from the representation

$$\lambda(r) = \max_{\substack{\|x\|_2=1\\ x>0}} \langle x, A_r x\rangle \tag{1.11}$$

we have, using (1.2) and (1.9),

$$\begin{aligned}&r\to\lambda(r) \text{ is strictly increasing with}\\ &\lim_{r\downarrow-\infty}\lambda(r) = 0,\ \lambda(0)<1,\ \lim_{r\uparrow\infty}\lambda(r) = \infty.\end{aligned} \tag{1.12}$$

Consequently we can define the following three objects:

(1.13) $$r^* \text{ is the unique solution of } \lambda(r) = 1,$$

(1.14) $$\theta^* = (\lambda'(r^*))^{-1},$$

(1.15) $$\nu^*(i,j) = \tau_{r^*}(i) A_{r^*}(i,j) \tau_{r^*}(j).$$

The solution of the variational problems (1.5) and (1.6) in Theorem 1 is now given as follows:

THEOREM 2A. *For every g satisfying* (1.1) *the minimum z and the minimizers $\bar{\theta}$ and $\bar{\nu}$ of* (1.5) *and* (1.6) *are*

(1.16) $$z = r^*,$$

(1.17) $$\bar{\theta} = \theta^*,$$

(1.18) $$\bar{\nu} = \nu^*.$$

THEOREM 2B. *Let g satisfy* (1.1). *Define g_β by*

(1.19) $$g_\beta(i) = \beta g(i), \qquad \beta \in (0, \infty).$$

Let $\theta^*(\beta)$ denote the quantity θ^* defined in (1.14) *when g is replaced by g_β. Then*

(1.20) $$\begin{gathered} \beta \to \theta^*(\beta) \text{ is analytic}, \\ \lim_{\beta \downarrow 0} \theta^*(\beta) = 0 \text{ and } \lim_{\beta \uparrow \infty} \theta^*(\beta) = 1. \end{gathered}$$

COROLLARY $*$. *Let $\lambda(\beta, r)$ denote the largest eigenvalue in* (1.10) *for g_β defined in* (1.19). *Then $\beta \to \theta^*(\beta)$ is strictly increasing iff*

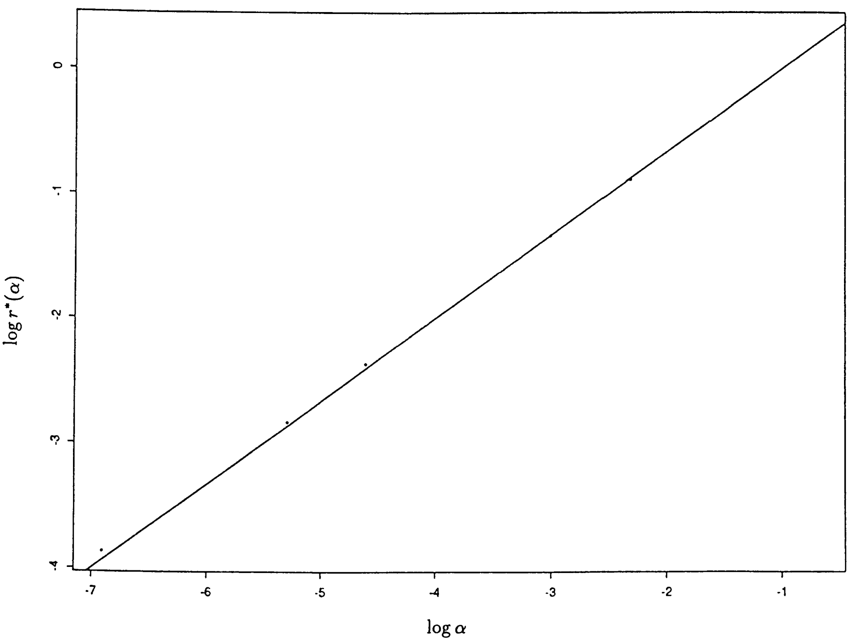
$(*)$ $$\left[\frac{\partial^2\lambda}{\partial r^2}\frac{\partial\lambda}{\partial\beta} - \frac{\partial^2\lambda}{\partial\beta\,\partial r}\frac{\partial\lambda}{\partial r}\right](\beta, r^*(\beta)) > 0.$$

In Section 3.6 we shall argue why $(*)$ is plausible at least when g, in addition to (1.1), has the property that $i \to g(i)/i$ is strictly increasing. Numerical evaluations of $r^*(\alpha)$ and $\theta^*(\alpha)$ are shown in Figure 1.

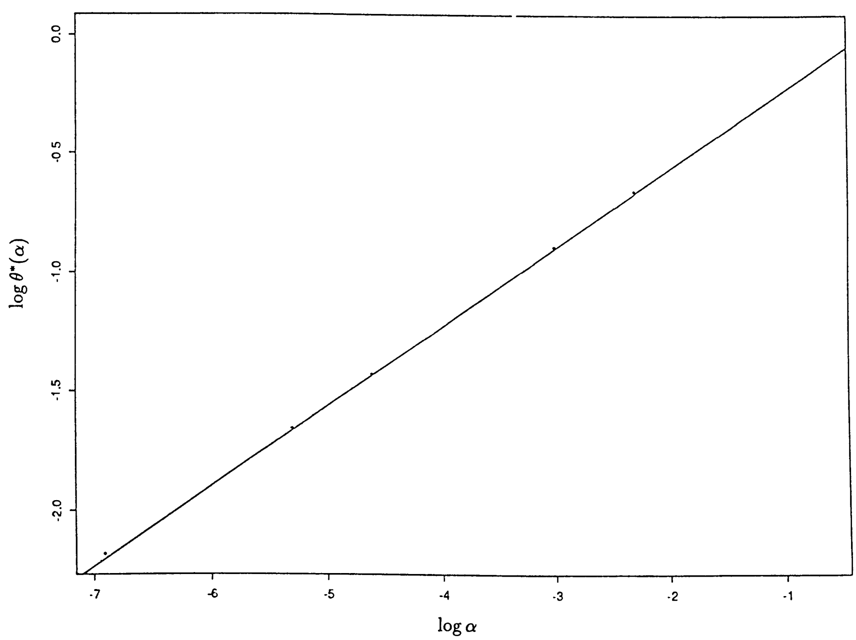
1.3. *Path properties.* The benefit of the variational approach to calculating the growth rate in (1.4) is that knowledge of the minimizers allows us to identify the asymptotic behavior of some functionals of the path, which we define below. This fact has been exploited in different contexts [see, e.g., Donsker and Varadhan (1975), Eisele and Lang (1987) and Baillon, Clément, Greven and den Hollander (1993)].

Fix g. Define Q_n^g by setting

(1.21) $$\frac{dQ_n^g}{dP_n} = \frac{1}{Z_n^g}\exp\left[-\sum_{x\in\mathbb{Z}} g(l_n(x))\right],$$



(a)



(b)

FIG. 1. *Numerical evaluation of* (a) $r^*(\alpha)$ *and* (b) $\theta^*(\alpha)$ *as a function of* α *for the model in* (0.1) *and* (0.2) [*i.e.*, $g(l) = \beta l^2$ *with* $\beta = -\log(1-\alpha)$] *plotted logarithmically. The computation is based on the representation* (1.2), (1.3), (1.9), (1.10), (1.13), (1.14) *with a* 100×100 *truncation of* A_r. *The dots are the numerical results. The straight lines have slope* 2/3 (*resp.* 1/3), *which suggests that* $r^*(\alpha) \sim C_1\alpha^{2/3}$ *and* $\theta^*(\alpha) \sim C_2\alpha^{1/3}$ *as* $\alpha \to 0$. *Note the remarkably good fit for* α *up to* 0.5. *The authors thank R. Gill for providing them with the simulations and the figures.*

where Z_n^g is the normalization factor [recall (0.1)]. Define the following functionals of the n-step path $(S_i)_{i=0}^n$:

$$\begin{aligned} M_n^+ &= \max_{0\le i\le n} S_i, \\ M_n^- &= \min_{0\le i\le n} S_i, \\ R_n &= (M_n^-, M_n^+); \end{aligned} \tag{1.22}$$

$$\begin{aligned} m_n^+(x) &= \sum_{0\le i<n} 1\{S_i = x, S_{i+1} = x+1\}, \\ m_n^-(x) &= \sum_{0\le i<n} 1\{S_i = x, S_{i+1} = x-1\}; \end{aligned} \tag{1.23}$$

$$\begin{aligned} \hat{\theta}_n &= \frac{1}{n} S_n, \\ \hat{\phi}_n &= \frac{1}{n}|R_n|; \end{aligned} \tag{1.24}$$

$$\begin{aligned} \hat{\nu}_n^+ &= |R_n|^{-1} \sum_{x\in R_n} \delta_{(m_n^+(x-1),\, m_n^+(x))}, \\ \hat{\nu}_n^- &= |R_n|^{-1} \sum_{x\in R_n} \delta_{(m_n^-(x),\, m_n^-(x-1))}. \end{aligned} \tag{1.25}$$

There is an obvious symmetry under Q_n^g between a path and its reflected image in 0. Nevertheless we have to distinguish between the path going right or going left:

THEOREM 3. *For every g satisfying* (1.1) *and for every neighborhood* $U(\theta^*)$ *of* θ^* *and* $U(\nu^*)$ *of* ν^* *(in the weak topology)*,

$$\begin{aligned} &\lim_{n\to\infty} Q_n^g\big(\hat{\theta}_n, \hat{\phi}_n \in U(\theta^*) | \hat{\theta}_n > 0\big) \\ &\quad = \lim_{n\to\infty} Q_n^g\big(-\hat{\theta}_n, \hat{\phi}_n \in U(\theta^*) | \hat{\theta}_n < 0\big) = 1, \end{aligned} \tag{1.26}$$

$$\begin{aligned} &\lim_{n\to\infty} Q_n^g\big(\hat{\nu}_n^+ \in U(\nu^*) | \hat{\theta}_n > 0\big) \\ &\quad = \lim_{n\to\infty} Q_n^g\big(\hat{\nu}_n^- \in U(\nu^*) | \hat{\theta}_n < 0\big) = 1. \end{aligned} \tag{1.27}$$

The relation (1.27) has an interesting consequence. Because

$$l_n(x) = m_n^+(x) + m_n^-(x), \tag{1.28}$$

$$\begin{aligned} m_n^-(x) &= m_n^+(x-1) - 1 \quad \text{for } x \in (0, S_n] \text{ and } S_n > 0, \\ m_n^+(x) &= m_n^-(x+1) - 1 \quad \text{for } x \in [S_n, 0) \text{ and } S_n < 0, \end{aligned} \tag{1.29}$$

it follows from (1.26) and (1.27) that the empirical distributions of local times, that is,

$$\hat{\mu}_n = |R_n|^{-1} \sum_{x\in R_n} \delta_{l_n(x)}, \quad n \in \mathbb{N}, \tag{1.30}$$

satisfy the following corollary.

COROLLARY. *For every neighborhood* $U(\mu^*)$ *of* μ^*,

$$\lim_{n\to\infty} Q_n^g(\hat{\mu}_n \in U(\mu^*)) = 1 \tag{1.31}$$

with μ^* *given by*

$$\mu^*(k) = \sum_{\substack{i,j\\ i+j-1=k}} \nu^*(i,j), \qquad k \in \mathbb{N}. \tag{1.32}$$

1.4. *Relation to earlier work.* In Greven and den Hollander (1992) and Baillon, Clément, Greven and den Hollander (1991) we have carried out the foregoing three-step program, but there the function g was in a different class, namely,

(1.33)(i) $g(0) = 0,$

(1.33)(ii) $g(i+j) < g(i) + g(j)$ for all $i, j \in \mathbb{N}$,

(1.33)(iii) $\lim_{i\to\infty} g(i)/i = 0,$

that is, (1.33)(ii) and (iii) are exactly the opposite of (1.1)(ii) and (iii). This corresponds to a self-attractive rather than a self-repellent random walk. The model giving rise to (1.33) is a branching random walk in random environment.

The proofs of Theorems 1–3 in Sections 2–4 make use of techniques developed in the preceding papers, but there are some significant changes in the results and consequently in the proofs. For instance, under (1.33) it turns out that for small θ the infimum over $\nu \in M_\theta$ in (1.6) is not attained in M_θ. Moreover, the infimum over $\theta \in (0,1]$ in (1.5) is attained at $\theta = 0$, resulting in $z = 0$.

2. Proof of Theorem 1. The proof of Theorem 1 uses ideas and results from Sections 2 and 3 in Greven and den Hollander (1992). There are four subsections and five lemmas. The main point is to show that if we condition the simple random walk to be at site $\lfloor \theta n \rfloor$ at time n, then its conditional law may be approximated on an exponential scale by the law of the random walk with drift θ, while the exponential functional in (0.8) may be approximated by a sum involving the total local times $l(x) = \lim_{n\to\infty} l_n(x)$ in the strip $[0, \lfloor \theta n \rfloor]$ ($\theta \in (0,1]$). The resulting expression is analyzed with the help of large deviation techniques for the empirical pair distribution of an underlying Markov chain.

Before we start, let us agree on some notation. Throughout the paper we assume that the function g satisfies (1.1). We also need a function $\delta\colon \mathbb{N} \to \mathbb{N}$ with the property

$$\begin{aligned} &\delta(n)/\log n \to \infty, \\ &\delta(n)/n \to 0, \qquad n \to \infty, \end{aligned} \tag{2.1}$$

which otherwise may be picked arbitrarily.

Define

$$\Omega = \{S = (S_i)_{i\in\mathbb{Z}} : S_0 = 0,\ S_i < 0 \text{ for } i < 0,\ |S_{i+1} - S_i| = 1 \text{ for } i \in \mathbb{Z}\}, \tag{2.2}$$

$$l(x, S) = \sum_{i\in\mathbb{Z}} 1\{S_i = x\}, \qquad S \in \Omega. \tag{2.3}$$

Let P_θ ($\theta \in (0,1]$) denote the law on Ω of the random walk with drift θ (i.e., the increments are i.i.d. $+1$ or -1 with probability $\frac{1}{2}(1+\theta)$ [resp. $\frac{1}{2}(1-\theta)$]. Think of this process as starting at $-\infty$ at time $-\infty$ and first hitting 0 at time 0 (or start the walk at 0 at time 0 and extend it to negative times by running a random walk with drift $-\theta$ from 0 conditioned on never returning to 0). Under the law P_θ the process $\{l(x,S)\}_{x\in\mathbb{Z}}$ is stationary and a.s. finite. Let E_θ denote expectation under P_θ.

Define

$$\Omega^n = \{S^n = (S_i)_{i=0}^n : S_0 = 0,\ |S_{i+1} - S_i| = 1 \text{ for } 0 \le i < n\}, \tag{2.4}$$

$$l(x, S^n) = \sum_{i=0}^{n} 1\{S_i = x\}, \qquad S^n \in \Omega^n. \tag{2.5}$$

Let P^n denote the law of simple random walk on Ω^n and let E^n denote expectation under P^n. Similarly define P_θ^n and E_θ^n for the random walk with drift θ.

Finally, define the strips

$$\begin{aligned} I_n(\theta) &= [0, \lfloor \theta n \rfloor], \\ J_n(\theta) &= [-\delta(n), \lfloor \theta n \rfloor + \delta(n)] \end{aligned} \tag{2.6}$$

and use the symbols $\asymp$, $\lesssim$ and $\gtrsim$ to denote (in)equality up to a subexponential factor as $n \to \infty$.

2.1. *Paths do not spill over on a linear scale.* Let

$$G(S^n) = \sum_x g(l(x, S^n)), \qquad S^n \in \Omega^n. \tag{2.7}$$

Then the expectation in (1.4) of Theorem 1 is

$$E^n(\exp[-G(S^n)]). \tag{2.8}$$

Let

$$A_n^1 = \{S^n \in \Omega_n : S_i \in (-\delta(n), S_n + \delta(n)) \text{ for } 0 < i \le n\}. \tag{2.9}$$

LEMMA 1.

$$E^n(\exp[-G(S^n)]) \asymp E^n(\exp[-G(S^n)]1\{S^n \in A_n^1\}). \tag{2.10}$$

PROOF. It suffices to prove that

$$\begin{aligned} &E^n(\exp[-G(S^n)]1\{S^n \in \Omega^n \setminus A_n^1\}) \\ &\quad \lesssim E^n(\exp[-G(S^n)]1\{S^n \in A_n^1\}). \end{aligned} \tag{2.11}$$

We shall do so by constructing maps

$$T_n : \Omega^n \setminus A_n^1 \to A_n^1 \tag{2.12}$$

with the following properties:

$$G(T_n(S^n)) \le G(S^n) \quad \text{for all } S^n \in \Omega^n \setminus A_n^1, \tag{2.13}$$

$$\left|\{S^n \in \Omega^n \setminus A_n^1 : T_n(S^n) = S'^n\}\right| \le (2n)^{n/2\delta(n)} \quad \text{for all } S'^n \in A_n^1. \tag{2.14}$$

This will give (2.11) via the estimates

$$\begin{aligned} &E^n\big(\exp[-G(S^n)]1\{S^n \in \Omega^n \setminus A_n^1\}\big)\\ &\quad\le E^n\big(\exp[-G(T_n(S^n))]1\{S^n \in \Omega^n \setminus A_n^1\}\big)\\ &\quad\le (2n)^{n/2\delta(n)} E^n\big(\exp[-G(S'^n)]1\{S'^n \in A_n^1\}\big). \end{aligned} \tag{2.15}$$

Note that T_n preserves probability because every $S^n \in \Omega^n$ has probability 2^{-n} under the law of simple random walk.

We finish the proof by exhibiting T_n through the following algorithm: First look for the leftmost point in the n-step path, say x, and its first hitting time $i(x)$. If $x \le -\delta(n)$, then reflect around x the piece of the path between time 0 and $i(x)$, and shift the whole path by $2x$ so that it again starts at 0. Repeat this procedure until the image path lies to the right of $-\delta(n)$. Next look for the rightmost point in the image path obtained so far, say y, and its last hitting time $i(y)$. Say z is its endpoint at time n. If $y \ge z + \delta(n)$, then reflect around y the piece of the path between time $i(y)$ and n. Repeat this procedure until the image path has its endpoint less than distance $\delta(n)$ away from its rightmost point.

Property (2.13) is an immediate consequence of (1.1)(ii): Each reflection lowers the value of G, because at some of the sites it splits off a part of the local time and moves this part to a site not yet visited [so (2.13) follows inductively]. Property (2.14) holds because the total number of reflections is at most $n/2\delta(n)$ (= length of path/minimum distance between reflection points) and because at most $2n$ sites can be reflection points ($= \|[-n, n)\|$). □

Note that the images under the map T_n of a path and its reflected image in 0 are the same. We have chosen to consider in the proof only those paths which move to the right eventually rather than to the left.

2.2. *Paths do not have zero speed.* Recall (2.1). Let

$$A_n^2 = \{S^n \in A_n^1 : S_n > \delta(n)\}. \tag{2.16}$$

LEMMA 2.

$$E^n\big(\exp[-G(S^n)]1\{S^n \in A_n^1\}\big) \simeq E^n\big(\exp[-G(S^n)]1\{S^n \in A_n^2\}\big). \tag{2.17}$$

PROOF. Each path in $A_n^1 \setminus A_n^2$ covers no more than $3\delta(n)$ sites. Hence, by (1.1)(ii),

$$G(S^n) \geq 3\delta(n) g\left(\left\lfloor \frac{n}{3\delta(n)} \right\rfloor\right), \qquad S^n \in A_n^1 \setminus A_n^2. \tag{2.18}$$

This lower bound grows faster than linear in n, by (1.1)(iii). □

2.3. *Conditioning on the speed and inserting total local times.* For $B \subset \mathbb{Z}$ finite and $k \in \mathbb{N}$, define

$$G(B,S) = \sum_{x \in B} g(l(x,S)), \qquad S \in \Omega, \tag{2.19}$$

$$A(B,k) = \left\{S \in \Omega \colon \sum_{x \in B} l(x,S) = k,\, l(x,S) = 1 \text{ for } x \in \partial B\right\}. \tag{2.20}$$

We shall need these quantities for $B = I_n(\theta)$, $J_n(\theta)$ and $k = n$, $n + 2\delta(n)$ [recall (2.1) and (2.6)]. Proposition 1 contains an assumption and a function $J(\theta)$ that will be verified (resp. identified) in Proposition 2 in Section 2.4.

PROPOSITION 1. *Suppose that there exists* $J\colon (0,1] \to \mathbb{R}^+$ *continuous with* $\lim_{\theta \downarrow 0} J(\theta) = \infty$ *such that*

$$\lim_{n\to\infty} \frac{1}{n} \log E_\theta\big(\exp[-G(I_n(\theta),S)] 1\{S \in A(I_n(\theta),n)\}\big) = -J(\theta) \tag{2.21}$$

and suppose that the same limit is obtained along any sequence $\theta_n \to \theta$. *Then*

$$\begin{aligned} &\lim_{n\to\infty} \frac{1}{n} \log E^n\big(\exp[-G(S^n)] 1\{S^n \in A_n^2\}\big) \\ &\quad = -\inf_{\theta \in (0,1]} [J(\theta) + I(\theta)], \end{aligned} \tag{2.22}$$

where

$$I(\theta) = \tfrac{1}{2}(1+\theta)\log(1+\theta) + \tfrac{1}{2}(1-\theta)\log(1-\theta). \tag{2.23}$$

PROOF. Let

$$A_n^2(\theta) = \{S^n \in A_n^2 \colon S_n = \lfloor \theta n \rfloor\} \tag{2.24}$$

and note that $A_n^2 = \bigcup_{\theta \in (\delta(n)/n, 1]} A_n^2(\theta)$. The proof proceeds in six steps.

The first step is to write

$$\begin{aligned} &E^n\big(\exp[-G(S^n)] 1\{S^n \in A_n^2\}\big) \\ &\quad = \int_{\theta \in (\delta(n)/n, 1]} \{d(\theta n) E^n\big(\exp[-G(S^n)] \mid S^n \in A_n^2(\theta)\big) \\ &\qquad\qquad P^n\big(S^n \in A_n^2(\theta)\big)\}. \end{aligned} \tag{2.25}$$

The second step is to observe that, for every $\theta' \in (0,1]$, under the law $P_{\theta'}^n$ all paths S^n ending at a given site x have the same probability, namely, $(\frac{1}{2}(1+\theta'))^{(n+x)/2}(\frac{1}{2}(1-\theta'))^{(n-x)/2}$. Therefore,

$$E^n(\exp[-G(S^n)] | S^n \in A_n^2(\theta)) = E_{\theta'}^n(\exp[-G(S^n)] | S^n \in A_n^2(\theta)) \tag{2.26}$$

for all θ and θ', that is, the conditional expectations are independent of the drift θ' of the random walk.

The third step is to pick $\theta' = \theta$ in (2.26) and substitute the resulting relation into (2.25) to get

$$\begin{aligned} &E^n(\exp[-G(S^n)]1\{S^n \in A_n^2\}) \\ &\quad = \int_{\theta \in (\delta(n)/n, 1]} \{d(\theta n) E_\theta^n(\exp[-G(S^n)]1\{S^n \in A_n^2(\theta)\}) \\ &\qquad \times [P^n(S^n \in A_n^2(\theta))/P_\theta^n(S^n \in A_n^2(\theta))]\}. \end{aligned} \tag{2.27}$$

The fourth step is the following lemma.

LEMMA 3. *Uniformly in θ outside any neighborhood of* 0,

$$\begin{aligned} &E_\theta^n(\exp[-G(S^n)]1\{S^n \in A_n^2(\theta)\}) \\ &\quad \simeq E_\theta(\exp[-G(I_n(\theta), S)]1\{S \in A(I_n(\theta), n)\}). \end{aligned} \tag{2.28}$$

We prove Lemma 3 later.

The fifth step is to note that

$$\begin{aligned} &P^n(S^n \in A_n^2(\theta))/P_\theta^n(S^n \in A_n^2(\theta)) \\ &\quad = \{(1+\theta)^{\frac{1}{2}(n+\lfloor \theta n \rfloor)}(1-\theta)^{\frac{1}{2}(n-\lfloor \theta n \rfloor)}\}^{-1} \\ &\quad \simeq \exp[-nI(\theta)] \quad \text{uniformly in } \theta, \end{aligned} \tag{2.29}$$

with $I(\theta)$ given by (2.23).

The sixth and final step is to substitute (2.28) into (2.27) and to apply Varadhan's theorem to the integral [see Deuschel and Stroock (1989), Theorem 2.1.10 and Exercise 2.1.20]. Via (2.21) and (2.29) this yields

$$\lim_{n\to\infty} \frac{1}{n} \log E^n(\exp[-G(S^n)]1\{S^n \in A_n^2\}) = \sup_{\theta \in (0,1]} [-J(\theta) - I(\theta)]. \tag{2.30}$$

The facts that θ in (2.21), (2.28) and (2.29) may be replaced by $\theta_n \to \theta$, that $J(\theta)$ and $I(\theta)$ are both continuous and that $\lim_{\theta \downarrow 0} J(\theta) = \infty$ are all needed to apply Varadhan's theorem. □

Thus it remains to prove Lemma 3. Note that the l.h.s. of (2.28) involves local times up to time n [recall (2.7)] and the r.h.s. total local times up to time ∞ [recall (2.19)].

PROOF OF LEMMA 3. In the sequel view S^n, Ω^n as the projection of S, Ω onto the coordinates $0, 1, \dots, n$. Let

$$A_n^3(\theta) = \{S \in \Omega: S^n \in A_n^2(\theta), S_i < -\delta(n) \text{ for all } i < -\delta(n), \\ S_i > \lfloor \theta n \rfloor + \delta(n) \text{ for all } i > n + \delta(n)\}, \tag{2.31}$$

that is, the path before time 0 and after time n moves straight through the boundary layer $J_n(\theta) \setminus I_n(\theta)$ and then stays outside of $J_n(\theta)$. Because $G(S^n)$ only depends on the path between time 0 and n, and because

$$P_\theta(S \in A_n^3(\theta))/P_\theta^n(S^n \in A_n^2(\theta)) = \theta\left(\tfrac{1}{2}(1 + \theta)\right)^{2\delta(n)} \tag{2.32}$$

[θ is the escape probability of random walk with drift θ; recall the remark below (2.3)], we have uniformly in θ outside any neighborhood of 0 [since $\delta(n) = o(n)$],

$$\text{l.h.s.}(2.28) \simeq E_\theta\left(\exp[-G(S^n)] 1\{S \in A_n^3(\theta)\}\right). \tag{2.33}$$

Now note that for all $S \in A_n^3(\theta)$,

$$\begin{aligned} l(x, S^n) &= l(x, S) && \text{for } x \in I_n(\theta) \\ &= l(x, S) - 1 && \text{for } x \in J_n(\theta) \setminus I_n(\theta) \\ &= 0 && \text{for } x \notin J_n(\theta), \end{aligned} \tag{2.34}$$

so that [recall (2.7) and (2.19)]

$$\begin{aligned} G(S^n) &= \sum_{x \in I_n(\theta)} g(l(x, S)) + \sum_{x \in J_n(\theta) \setminus I_n(\theta)} g(l(x, S) - 1) \\ &\leq G(J_n(\theta), S). \end{aligned} \tag{2.35}$$

Upper bound: Note that $A_n^3(\theta) \subseteq A(J_n(\theta), n + 2\delta(n))$ and hence from (2.35),

$$\text{r.h.s.}(2.33) \leq E_\theta\left(\exp[-G(J_n(\theta), S)] 1\{S \in A(J_n(\theta), n + 2\delta(n))\}\right). \tag{2.36}$$

Now, the r.h.s. of (2.28) and (2.36) look identical, the only difference being that $\lfloor \theta n \rfloor$ and n are perturbed to become $\lfloor \theta n \rfloor + 2\delta(n)$ and $n + 2\delta(n)$ [use the stationarity of the process $\{l(x, S)\}_{x \in \mathbb{Z}}$ under the law P_θ to shift $J_n(\theta)$ by $\delta(n)$ to get the set $[0, \lfloor \theta n \rfloor + 2\delta(n)]$, which is a perturbation of $I_n(\theta)$]. However, because $\delta(n) = o(n)$, and because the supposition in Lemma 3 states that the same limit is obtained in (2.21) along any sequence $\theta_n \to \theta$, it follows via (2.33) and (2.36) that

$$\text{l.h.s.}(2.28) \lesssim \text{r.h.s.}(2.28). \tag{2.37}$$

Lower bound: Note that $A_n^2(\theta)$ contains the projection of $A(I_n(\theta), n)$ on the coordinates $0, 1, \dots, n$ and that for all $S \in A(I_n(\theta), n)$,

$$\begin{aligned} l(x, S^n) &= l(x, S) && \text{for } x \in I_n(\theta) \\ &= 0 && \text{for } x \notin I_n(\theta), \end{aligned} \tag{2.38}$$

so that

$$G(S^n) = G(I_n(\theta), S). \tag{2.39}$$

Hence it trivially follows that

$$\text{l.h.s.}(2.28) \geq \text{r.h.s.}(2.28). \tag{2.40}$$

□

2.4. *Large deviations for total local times.*

PROPOSITION 2. *The suppositions in Proposition* 1 *hold with J given by*

$$J(\theta) = \theta \inf_{\nu \in M_\theta} [\langle g \circ a, \nu \rangle + I_\theta(\nu)], \tag{2.41}$$

where

$$I_\theta(\nu) = \sum_{i,j} \nu(i,j) \log\left(\frac{\nu(i,j)}{\hat{\nu}(i) P_\theta(i,j)}\right), \tag{2.42}$$

$$P_\theta(i,j) = \binom{i+j-2}{i-1} [\tfrac{1}{2}(1+\theta)]^i [\tfrac{1}{2}(1-\theta)]^{j-1}, \tag{2.43}$$

M_θ *is as defined in* (1.8), $a(i,j) = i + j - 1$ *and* $\hat{\nu}(i) = \sum_j \nu(i,j)$.

PROOF. Fix $\theta \in (0, 1]$. Let us first recall the quantity in (2.21):

$$E_\theta\left(\exp\left[-\sum_{x=0}^{\lfloor \theta n \rfloor} g(l(x))\right] 1\left\{\sum_{x=0}^{\lfloor \theta n \rfloor} l(x) = n, l(0) = l(\lfloor \theta n \rfloor) = 1\right\}\right). \tag{2.44}$$

Here we drop S from the notation and write $l(x)$ instead of $l(x, S)$ [recall (2.19) and (2.20)]. In Section 3 of Greven and den Hollander (1992) (GH) we computed the exponential growth rate of exactly the same quantity, except that there our function g was in the class (1.33) instead of (1.1) (see Section 1.4). We shall now have to see how our techniques may be carried over. There are some significant changes. Before we start with the formal proof, we first sketch the main line of thought in order to give the reader some guidance.

Define the empirical distributions

$$\mu_N = \frac{1}{N} \sum_{x=0}^{N-1} \delta_{l(x)}, \qquad N \in \mathbb{N}, \tag{2.45}$$

on $\mathscr{P}(\mathbb{N})$, the set of probability measures on $\mathbb{N}$. Then (2.44) may be rewritten as

$$E_\theta\left(\exp[-K_n \langle g, \mu_{K_n} \rangle] 1\{\langle 1, \mu_{K_n} \rangle = L_n, l(0) = l(K_n - 1) = 1\}\right) \tag{2.46}$$

with $K_n = \lfloor \theta n \rfloor + 1$, $L_n = n/K_n$ and $1\colon \mathbb{N} \to \{1\}$. Under the law P_θ the process $\{l(x)\}_{x \geq 0}$ is stationary mixing and, therefore, one expects that $(\mu_N)_{N \in \mathbb{N}}$ satisfies the large deviation principle on $\mathscr{P}(\mathbb{N})$ (in the weak topology) with some rate function, say $I(\mu)$. Because $K_n \sim \theta n$ and $L_n \to \theta^{-1}$, one thus expects that

$$\lim_{n\to\infty}\frac{1}{n}\log(2.46) = \theta \lim_{K_n\to\infty}\frac{1}{K_n}\log(2.46) = \theta \sup_{\mu\in N_\theta}\left[-\langle g,\mu\rangle - I(\mu)\right] \tag{2.47}$$

with

$$N_\theta = \left\{\mu \in \mathscr{P}(\mathbb{N}) : \langle 1,\mu\rangle = \theta^{-1}\right\}. \tag{2.48}$$

Now, we shall see that (2.47) and (2.48) are basically correct. However, there are some problems along the way:

1. How to identify the rate function $\mu \to I(\mu)$? The process $\{l(x)\}_{x\ge 0}$ is not Markovian.
2. $\mu \to \langle g,\mu\rangle$ is not continuous on $\mathscr{P}(\mathbb{N})$ (in the weak topology).
3. The indicator set in (2.46) depends on n and is not closed (in the weak topology).

We conclude this section by explaining how to resolve these three problems:

Problem 1. The main observation in GH to handle Problem 1 is that $\{l(x)\}_{x\ge 0}$ is a two-block functional of a Markov process, namely,

$$\begin{aligned} l(x) &= m^+(x) + m^-(x),\\ m^+(x) &= \sum_{i\ge 0} 1\{S_i = x,\, S_{i+1} = x+1\},\\ m^-(x) &= \sum_{i\ge 0} 1\{S_i = x,\, S_{i+1} = x-1\}; \end{aligned} \tag{2.49}$$

$$m^-(x) = m^+(x-1) - 1\{x\ge 1\} \qquad P_\theta\text{-a.s.}; \tag{2.50}$$

(2.51) $\{m^+(x)\}_{x\ge 0}$ is Markovian with transition matrix $P_\theta(i,j)$ given by (2.43)

[see GH, Section 3a]. This allows us to introduce the *empirical pair distributions*

$$\nu_N = \frac{1}{N}\sum_{x=1}^{N}\delta_{(m^+(x-1),\, m^+(x))}, \qquad N\in\mathbb{N}, \tag{2.52}$$

and to rewrite (2.46) as

$$E_\theta\Big(\exp\left[-K_n\langle g\circ a, \nu_{K_n}\rangle\right]1\big\{\langle a,\nu_{K_n}\rangle = L_n,\, m^+(-1) = m^+(0) = m^+(K_n-2) = m^+(K_n-1) = 1\big\}\Big). \tag{2.53}$$

Now suppose that $(\nu_N)_{N\in\mathbb{N}}$ satisfies the large deviation principle on $\mathscr{P}(\mathbb{N}^2)$. Then its rate function $I_\theta(\nu)$ has the standard form for Markov processes given by (2.42) [see Ellis (1985), page 19], and hence one guesses that

$$\lim_{n\to\infty}\frac{1}{n}\log(2.53) = \theta\sup_{\nu\in M_\theta}\left[-\langle g\circ a,\nu\rangle - I_\theta(\nu)\right] \tag{2.54}$$

with

$$M_\theta = \left\{\nu \in \mathscr{P}(\mathbb{N}^2) \colon \langle a, \nu\rangle = \theta^{-1},\ \sum_j \nu(i,j) = \sum_j \nu(j,i) \text{ for all } i\right\}. \tag{2.55}$$

[This is of course still subject to Problems 2 and 3 (in terms of ν).] Recalling that (2.53) is a reformulation of the expectation in (2.21), we see that the r.h.s. of (2.54) identifies J as the expression in (2.41). Thus, to solve Problem 1 we are left with establishing the large deviation principle for the process $\{m^+(x)\}_{x \in \mathbb{Z}}$ and its empirical pair distributions $(\nu_N)_{N \in \mathbb{N}}$.

In GH, Section 3a, it is pointed out that the Markov process with transition matrix P_θ given in (2.43) and (2.51) is not in the class of uniformly ergodic Markov processes for which large deviation principles have been derived in the literature. Therefore, some work is needed to handle the infinite state space $\mathbb{N}^2$. This obstacle is removed together with Problem 2.

Problem 2. This is dealt with in GH, Section 3c, by the method of truncation. Here it is shown that we may insert into (2.53) the indicator $1\{\sup_{0 \le x \le K_n} m_n^+(x) \le R\}$, compute the exponential growth rate as $n \to \infty$ and let $R \to \infty$ afterward. The proof depends on several estimates, all of which carry over but one, which we now show how to modify. For finite R the large deviation principle follows easily, so truncation also removes the obstacle just mentioned in proving (2.54).

The reader is asked at this point to look up GH, Section 3c, or to skip the argument. Replace GH (3.23) by

$$\begin{aligned} &G(X) - G(\tilde{T}(X)) \ge 0,\\ &G(X) = \sum_{k=1}^{K_n} g(X_{k-1} + X_k - 1),\\ &G(\tilde{T}X) = \sum_{k=1}^{K_n} g\big((\tilde{T}X)_{k-1} + (\tilde{T}X)_k - 1\big). \end{aligned} \tag{2.56}$$

Here $X = (X_k)_{k=0}^{K_n}$ are i.i.d. geometrically distributed $\mathbb{N}$-valued random variables with boundary condition $X_0 = X_{K_n} = 1$, and $\tilde{T}$ is a map that truncates the X_k exceeding R and adds the excess to some $X_{k'}$ sufficiently below R. The map $\tilde{T}$ is a slight modification of the map T that appears in GH, section 3c:

(i) Redefine $B^1 = \{k\colon X_k \le v(t-1),\ k \notin \delta A^2\}$.

(ii) From each $k \in A^2$ remove as many piles of size s such that at least $v(t-1)$ objects remain.

(iii) If $v \to \infty$, then the estimate $|B^2| \sim \lfloor \theta n \rfloor$ $(n \to \infty)$ carries over.

The inequality in (2.56) follows easily from the property $g(i) + g(j+\Delta) > g(i+\Delta) + g(j)$ for all $i < j$ and $\Delta > 0$, which is a consequence of (1.1)(iii). Here Δ is the excess that is moved to lower levels by the map $\tilde{T}$. If $v/u \to \infty$, then the estimate below GH (3.26) still holds. This solves Problem 2. [A more adapted proof can be based on (1.1)(iii). For reasons of space we have appealed to the setup in GH.]

Problem 3. This is dealt with in GH, Section 3d, by the method of perturbation. Here it is shown that we may thicken up the restriction $\langle a, \nu_{K_n}\rangle = L_n$ in (2.53) to a slab $\langle a, \nu_{K_n}\rangle \in [\theta^{-1} - \varepsilon, \theta^{-1} + \varepsilon]$ (recall that $L_n \to \theta^{-1}$), compute the exponential growth rate as $n \to \infty$ and let $\varepsilon \downarrow 0$ afterward. The proof again depends on several estimates, all of which carry over but one.

Replace GH (3.31) by

$$\begin{aligned} &0 \le G(X) - G(X') \le 2\varepsilon(n) g(2R-1) \\ &\text{when } \sum_{k=1}^{K_n} [X_k - X_k'] = \varepsilon(n) \text{ and } X_k' \ge X_k \ge R, \end{aligned} \tag{2.57}$$

where R is the truncation level [also see the remark before GH (3.34)] and $\varepsilon(n) \le \varepsilon n$.

The final step in the chain of arguments is explained in GH, Section 3e. For $R < \infty$ and $\varepsilon > 0$ the large deviation problem is standard and we end up with a variational formula similar to the the r.h.s. of (2.54), namely,

$$\theta \sup_{\nu \in M_\theta^{\varepsilon, R}} [-\langle g \circ a, \nu\rangle - I_\theta(\nu)] \tag{2.58}$$

with $M_\theta^{\varepsilon, R}$ the truncated perturbed set

$$\begin{aligned} M_\theta^{\varepsilon, R} = \Big\{ &\nu \in \mathscr{P}([1, R]^2) : \langle a, \nu\rangle \in [\theta^{-1} - \varepsilon, \theta^{-1} + \varepsilon], \\ &\sum_j \nu(i,j) = \sum_j \nu(j,i) \text{ for all } i \Big\}. \end{aligned} \tag{2.59}$$

It now only remains to show that (2.58) reduces to the r.h.s. of (2.41) as $\varepsilon \downarrow 0$ and $R \to \infty$ (in this order). This goes as follows. Put $\phi(\nu) = \langle g \circ a, \nu\rangle + I_\theta(\nu)$. First let $\varepsilon \to 0$. Because $\nu \to \phi(\nu)$ is continuous on $\bigcup_{\varepsilon > 0} \mathrm{M}_\theta^{\varepsilon, R}$, the supremum in (2.58) reduces to $\nu \in M_\theta^R = \bigcap_{\varepsilon > 0} M_\theta^{\varepsilon, R}$. Next let $R \to \infty$. We show that

$$\sup_{R < \infty} \sup_{\nu \in M_\theta^R} [-\phi(\nu)] = \sup_{\substack{\nu \in M_\theta \\ \langle g \circ a, \nu\rangle < \infty}} [-\phi(\nu)]. \tag{2.60}$$

Indeed, the supremum over ν may be restricted to the symmetric measures, that is, $\nu(i,j) = \nu(j,i)$ [see (3.29)]. Because $\nu \to I_\theta(\nu)$ is continuous on M_θ (see GH, Section 3e), it suffices to show that for every $\nu \in M_\theta$, symmetric with $\langle g \circ a, \nu\rangle < \infty$, there exists a sequence (ν_R), with $\nu_R \in M_\theta^R$, such that $\nu_R \to \nu$ weakly and

$$\limsup_{R \to \infty} \langle g \circ a, \nu_R\rangle = \langle g \circ a, \nu\rangle. \tag{2.61}$$

Here is an example of such a sequence. Pick (k,l) with $k \le l < R$ and $\nu(k,l) = \nu(l,k) > 0$, and put

$$\begin{aligned} &(2.62) \quad && \nu_R(i,j) = \nu_R(j,i), \\ & && \nu_R(i,j) = \nu(i,j) \qquad \text{if } i,j < R, (i,j) \ne (k,l), (k+1,l), \\ & && \nu_R(i,R) = \sum_{j \ge R} \nu(i,j) \quad \text{if } i < R, \\ &(2.63) && \nu_R(R,R) = \sum_{i,j>R} \nu(i,j), \\ & && \nu_R(k,l) = \nu(k,l) - \varepsilon_R, \\ & && \nu_R(k+1,l) = \nu(k+1,l) + \varepsilon_R, \end{aligned}$$

with ε_R chosen such that $\langle a, \nu_R \rangle = \langle a, \nu \rangle = \theta^{-1}$. Because of $\langle g \circ a, \nu \rangle < \infty$ and $g(i+j-1) \ge c(i+j-1)$ with $c = g(1) > 0$ [see (1.1)], it follows that $\varepsilon_R \to 0$ as $R \to \infty$. Because $R \to \nu_R$ is stochastically increasing, except at the points (k,l), $(k+1,l)$, (l,k) and $(l,k+1)$, we have (2.61) by monotone convergence. Finally, we remove the restriction from the r.h.s. of (2.60) by noting that $I_\theta(\nu) \ge 0$. We have therefore solved Problem 3.

This completes the proof of (2.21) in Proposition 1 and identifies J as given by (2.41). The remaining suppositions in Proposition 1 are easily settled. Indeed, the fact that everywhere θ may be replaced by $\theta_n \to \theta$ is a corollary of the previously sketched perturbation argument (see the end of GH, Section 3e), which also implies that $\theta \to J(\theta)$ is continuous on $(0,1]$. To get $\lim_{\theta \downarrow 0} J(\theta) = \infty$, note that for every $\nu \in M_\theta$ by (1.1)(ii),

$$(2.64) \qquad \begin{aligned} \langle g \circ a, \nu \rangle &= \sum_{i,j} g(i+j-1)\nu(i,j) \ge g\Big(\sum_{i,j} (i+j-1)\nu(i,j) \Big) \\ &= g(\theta^{-1}) \end{aligned}$$

and hence $J(\theta) \ge \theta g(\theta^{-1})$ from (2.41), because $I_\theta(\nu) \ge 0$. Let $\theta \downarrow 0$ and use (1.1)(iii). □

We can now collect Propositions 1 and 2 and Lemmas 1–3 and prove Theorem 1.

PROOF OF THEOREM 1. The expectation in (1.4) of Theorem 1 equals (2.8). Combine (2.10), (2.17), (2.22), (2.23) and (2.41) to get for the quantity z, defined in (1.4),

$$(2.65) \qquad z = \inf_{\theta \in (0,1]} \Big\{ \theta \inf_{\nu \in M_\theta} [\langle g \circ a, \nu \rangle + I_\theta(\nu)] + I(\theta) \Big\}.$$

Now note from (1.8) that $\nu \in M_\theta$ implies $\sum_{i,j} i\nu(i,j) = (1+\theta)/2\theta$ and $\sum_{i,j}(j-1)\nu(i,j) = (1-\theta)/2\theta$. Compare (1.3) with (2.43) to see that it follows from (2.42), that

$$
\begin{aligned}
I_\theta(\nu) = {}& \sum_{i,j} \nu(i,j)\log\left(\frac{\nu(i,j)}{\hat{\nu}(i)P(i,j)}\right) \\
& - \frac{1+\theta}{2\theta}\log(1+\theta) - \frac{1-\theta}{2\theta}\log(1-\theta).
\end{aligned}
\tag{2.66}
$$

The last two terms are precisely $-I(\theta)/\theta$, by (2.23), so after cancellation in (2.65) we end up with (1.5) and (1.6). [Recall (1.2) and $\langle g \circ a, \nu\rangle = \sum_{i,j} g(i+j-1)\nu(i,j)$.] □

3. Proof of Theorem 2. In this section we solve the variational problems (1.5) and (1.6). The reader is asked to recall the notation introduced in (1.2), (1.3) and (1.5)–(1.8). Along the way we shall be able to use some of the results on the solution of the variational problem in Baillon, Clément, Greven and den Hollander (1991) (BCGH), which is of exactly the same form but assumes that g is in the class (1.33) rather than in the class (1.1) we are now dealing with (recall Section 1.4).

3.1. *Existence of a minimizer in* (1.6).

PROPOSITION 3. *For every $\theta \in (0,1]$ there exists at least one $\bar{\nu} \in M_\theta$ such that*

$$\phi(\bar{\nu}) = K(\theta). \tag{3.1}$$

PROOF. Fix $\theta \in (0,1]$. The first observation is that the functional ϕ in (1.7) is lower semicontinuous on the set M_θ equipped with the weak topology. This fact uses only the property $g \geq 0$ and for a proof we can, therefore, refer to BCGH, Section 2.1, Lemma 2. A lower semicontinuous function on a compact set always attains a minimum. However, M_θ is not compact in the weak topology. The way out of this problem is the following procedure.

Define

$$M_{\theta,N} = \left\{\nu \in M_\theta: \sum_{i,j} g(i+j-1)\nu(i,j) \leq N\right\}, \qquad N \in \mathbb{R}^+. \tag{3.2}$$

We shall prove the following lemma.

LEMMA 4. *For N sufficiently large, $M_{\theta,N}$ is nonempty and*

$$\inf_{\nu \in M_{\theta,N}} \phi(\nu) < \inf_{\nu \in M_\theta \setminus M_{\theta,N}} \phi(\nu), \tag{3.3}$$

$$M_{\theta,N} \textit{ is compact in the weak topology}. \tag{3.4}$$

Lemma 4 implies Proposition 3, using the remarks preceding (3.2). □

PROOF OF LEMMA 4. To prove (3.3) we use (1.2) and rewrite $\phi(\nu)$ as

$$\phi(\nu) = \sum_{i,j} g(i+j-1)\nu(i,j) + I(\nu) \tag{3.5}$$

with

$$I(\nu) = \sum_{i,j} \nu(i,j)\log\left(\frac{\nu(i,j)}{\hat{\nu}(i)P(i,j)}\right). \tag{3.6}$$

Because P is a Markov matrix [recall (1.3)], $I(\nu)$ is a relative entropy and hence $I(\nu) \geq 0$. Therefore, $\phi(\nu) \geq \sum_{i,j} g(i+j-1)\nu(i,j)$, and so (3.3) will follow once we show that there exists at least one $\nu_\theta \in M_\theta$ with the properties

$$\begin{gathered} \phi(\nu_\theta) < \infty, \\ \sum_{i,j} g(i+j-1)\nu_\theta(i,j) < \infty. \end{gathered} \tag{3.7}$$

But this is easy; namely, pick

$$\begin{aligned} &\nu_\theta = a\delta_{(k,k)} + b\delta_{(k+1,k+1)} \text{ with } k = \left\lfloor \frac{1}{2\theta} \right\rfloor \text{ or } \left\lceil \frac{1}{2\theta} \right\rceil \text{ and } \textit{with } a, b \geq 0 \\ &\text{solving } a+b=1,\ a(2k-1)+b(2k+1) = \frac{1}{\theta}. \end{aligned} \tag{3.8}$$

To prove (3.4), first note that (1.1)(iii), together with Lebesgue's dominated convergence theorem, implies

$$\left.\begin{aligned} &(\nu_n) \in M_{\theta,N} \\ &\nu_n \to \nu \text{ weakly } (n \to \infty) \end{aligned}\right\} \quad \Rightarrow \quad \nu \in M_\theta. \tag{3.9}$$

Next use that $g \geq 0$. Hence Fatou's lemma gives

$$\liminf_{n\to\infty} \sum_{i,j} g(i+j-1)\nu_n(i,j) \geq \sum_{i,j} g(i+j-1)\nu(i,j), \tag{3.10}$$

and so in fact $\nu \in M_{\theta,N}$. Consequently, $M_{\theta,N}$ is closed. Finally note that $M_\theta \subseteq \{\nu \in \mathscr{P}(\mathbb{N}^2): \sum_{i,j}(i+j-1)\nu(i,j) \leq \theta^{-1}\}$. The latter set is compact by the Helly–Bray theorem (see, e.g., Breiman, Theorem 8.6 and Proposition 8.10). Consequently, $M_{\theta,N}$ is compact. □

3.2. *Variation over* ν: *Identification of the minimizer in* (1.6). In this section we shall identify the minimizer of (1.6) and compute the minimum. In particular, we shall see that the minimizer is unique. The result is stated in Proposition 4. We begin by introducing some quantities.

We approach the problem by the technique of variations. Given any minimizer $\bar{\nu}$, we can analyze the behavior of ϕ in a small neighborhood of $\bar{\nu}$. From the fact that ϕ is minimal at $\bar{\nu}$ we can deduce certain relations for $\bar{\nu}$, and by considering enough variations we should be able to actually determine what $\bar{\nu}$ looks like.

Formally, consider variations $\bar{\nu} + t\partial$, $t > 0$, around a minimizer $\bar{\nu}$ of (1.6) such that for t sufficiently small $\bar{\nu} + t\partial$ is still in the domain M_θ. The latter is ensured by the following restrictions on ∂:

$$
\begin{aligned}
\partial(i,j) &\geq 0 \quad \text{for all } i,j \text{ with } \bar{\nu}(i,j)=0,\\
\partial(i,j) &= 0 \quad \text{except at finitely many points,} \qquad (3.11)\\
\sum_j \partial(i,j) &= \sum_j \partial(j,i) \quad \text{for all } i,
\end{aligned}
$$

$$
\sum_{i,j} \partial(i,j) = 0, \qquad (3.12)
$$

$$
\sum_{i,j} (i+j-1)\partial(i,j) = 0. \qquad (3.13)
$$

Variations fulfilling (3.11)–(3.13) will be called admissible.

From the fact that ϕ is convex (see BCGH, Section 2.1, Lemma 2) and is minimal at $\bar{\nu}$ it follows that the following limit exists and is positive:

$$
\lim_{t \downarrow 0} \frac{1}{t}\left[\phi(\bar{\nu} + t\partial) - \phi(\bar{\nu})\right] \geq 0. \qquad (3.14)
$$

Using (3.14) in combination with (3.11)–(3.13) we shall be able to establish the following lemma.

LEMMA 5. *For every minimizer* $\bar{\nu} \in M_\theta$:

(i) $\bar{\nu}(i,j) = \bar{\nu}(j,i)$ *for all* i,j.
(ii) $\bar{\nu}(i,j) > 0$ *for all* i,j *if* $\theta < 1$.
(iii) *If* $\bar{\xi}, \bar{\eta}$ *are defined by*

$$
\begin{aligned}
\bar{\xi}(i,j) &= \log\left(\frac{\bar{\nu}(i,j)}{\hat{\bar{\nu}}(i)A(i,j)}\right),\\
\bar{\eta}(i,j) &= \frac{1}{2}\left[\bar{\xi}(i,j) + \bar{\xi}(j,i)\right],
\end{aligned} \qquad (3.15)
$$

then there exists $r \in \mathbb{R}$ *such that*

$$
\bar{\eta}(k,l) - \bar{\eta}(1,1) = r(k+l-2) \quad \textit{for all } k,l. \qquad (3.16)
$$

In order not to interrupt the flow of the argument, we defer the proof of Lemma 5 to Section 3.3 and continue by pointing out what (3.15) and (3.16) do for us.

PROPOSITION 4. *For every* $\theta \in (0,1)$ *the minimizer* $\bar{\nu} = \bar{\nu}(\theta) \in M_\theta$ *of* (1.6) *is unique and is given by*

$$\bar{\nu}(i,j) = \frac{1}{\lambda(r)} \tau_r(i) A_r(i,j) \tau_r(j), \tag{3.17}$$

$$r = r(\theta) \text{ is the unique solution of } \theta^{-1} = \frac{\lambda'(r)}{\lambda(r)}. \tag{3.18}$$

Here A_r *is defined in* (1.9) *and* $(\lambda(r), \tau_r)$ *in* (1.10). *The minimum is*

$$K(\theta) = \phi(\bar{\nu}) = \frac{r}{\theta} + \log \frac{1}{\lambda(r)}. \tag{3.19}$$

PROOF. First substitute (3.15) into (3.16) to obtain the relation

$$\begin{gathered} \bar{\nu}(i,j)\bar{\nu}(j,i) = R_r^2 \hat{\bar{\nu}}(i) \hat{\bar{\nu}}(j) A_r(i,j) A_r(j,i) \\ \text{with } R_r = \bar{\nu}(1,1) / \big(\hat{\bar{\nu}}(1) A_r(1,1)\big). \end{gathered} \tag{3.20}$$

Here note that the parameter r of (3.16) is merged with the matrix A to form A_r of (1.9). Because $\bar{\nu}$ is symmetric according to Lemma 5(i) and A_r is symmetric, this relation can be rewritten in the shape of a quadratic form, namely

$$\bar{\nu}(i,j) = R_r \big[\hat{\bar{\nu}}(i)\big]^{1/2} A_r(i,j) \big[\hat{\bar{\nu}}(j)\big]^{1/2}. \tag{3.21}$$

Next sum over j and use the fact that $\hat{\bar{\nu}} > 0$ according to Lemma 5(ii) to obtain

$$R_r^{-1} \big[\hat{\bar{\nu}}(i)\big]^{1/2} = \sum_j A_r(i,j) \big[\hat{\bar{\nu}}(j)\big]^{1/2}. \tag{3.22}$$

In words, $\hat{\bar{\nu}}^{1/2}$ is a strictly positive eigenvector of A_r with eigenvalue R_r^{-1}. Note that $\hat{\bar{\nu}}^{1/2} \in l^2(\mathbb{N})$ with $\|\hat{\bar{\nu}}^{1/2}\|_2 = 1$, because $\hat{\bar{\nu}} \in \mathscr{P}(\mathbb{N})$.

The next step is to use that A_r: $l^2(\mathbb{N}) \to l^2(\mathbb{N})$ is a strictly positive, self-adjoint and compact operator [recall the remark below (1.9); compactness follows from (1.1)(iii)]. This allows us to conclude the following:

$$\begin{aligned} &\big(R_r^{-1}, \hat{\bar{\nu}}^{1/2}\big) = (\lambda(r), \tau_r), \\ &\text{where } \lambda(r) \text{ is the largest eigenvalue of } A_r \\ &\text{and } \tau_r \text{ the corresponding eigenvector} \\ &\text{with } \|\tau_r\|_2 = 1,\ \tau_r > 0. \end{aligned} \tag{3.23}$$

The positivity, symmetry and compactness of A_r imply that $\lambda(r)$ is simple and that $r \to \lambda(r)$ and $r \to \tau_r$ are analytic (see BCGH, Section 2.5, Proposition 5).

It is now easy to determine the parameter r appearing in (3.16). Namely, because $\bar{\nu} \in M_\theta$, we must have

$$\begin{aligned} \theta^{-1} &= \sum_{i,j} (i+j-1)\bar{\nu}(i,j) \\ &= \frac{1}{\lambda(r)} \sum_{i,j} (i+j-1)\tau_r(i) A_r(i,j)\tau_r(j) \\ &= \frac{1}{\lambda(r)} \left\langle \tau_r, \left(\frac{\partial}{\partial r} A_r\right)\tau_r \right\rangle. \end{aligned} \tag{3.24}$$

On the other hand, because $\lambda(r) = \langle \tau_r, A_r \tau_r \rangle$, we have

$$\begin{aligned} \lambda'(r) &= \frac{\partial}{\partial r} \langle \tau_r, A_r \tau_r \rangle \\ &= \left\langle \tau_r, \left(\frac{\partial}{\partial r} A_r\right)\tau_r \right\rangle + \left\langle \left(\frac{\partial}{\partial r}\tau_r\right), A_r\tau_r \right\rangle + \left\langle \tau_r, A_r\left(\frac{\partial}{\partial r}\tau_r\right) \right\rangle. \end{aligned} \tag{3.25}$$

The last two terms are identical (because A_r is symmetric) and are equal to

$$\lambda(r) \left\langle \left(\frac{\partial}{\partial r}\tau_r\right), \tau_r \right\rangle = \frac{1}{2}\lambda(r) \frac{\partial}{\partial r} \langle \tau_r, \tau_r \rangle = 0, \tag{3.26}$$

where the last equality uses that $\|\tau_r\|_2 = 1$ for all r. Hence, combining (3.24)–(3.26), we have

$$\theta^{-1} = \frac{\lambda'(r)}{\lambda(r)}. \tag{3.27}$$

This equation has a unique solution $r = r(\theta)$ for every $\theta \in (0,1)$, because $r \to \lambda'(r)/\lambda(r)$ is strictly increasing and has limits 1 as $r \downarrow -\infty$ and ∞ as $r \uparrow \infty$. The former property is the same as saying that $r \to \lambda(r)$ is strictly log-convex, which is proved in BCGH, Section 3.3, Lemma 12, by using a theorem of Kato. The latter property is an easy consequence of the form of A_r. This completes the proof of the first half of Proposition 4, namely (3.17) and (3.18).

To get (3.19), the second half of Proposition 4, simply substitute (3.17) into (1.7) to compute, recalling (1.8) and (1.9) and using $\bar{\nu} \in M_\theta$,

$$\begin{aligned} K(\theta) = \phi(\bar{\nu}) &= \sum_{i,j} \bar{\nu}(i,j) \log\left(\frac{\bar{\nu}(i,j)}{\hat{\bar{\nu}}(i) A(i,j)}\right) \\ &= \log\frac{1}{\lambda(r)} + \sum_{i,j} \bar{\nu}(i,j)\log\left(\frac{A_r(i,j)}{A(i,j)}\right) \\ &= \log\frac{1}{\lambda(r)} + \frac{r}{\theta}. \end{aligned} \tag{3.28}$$

□

3.3. *Proof of Lemma* 5. (i) Define the measure ν^S by $\nu^S(i,j) = \nu(j,i)$. Because ϕ is convex and $\phi(\nu) = \phi(\nu^S)$, and because $\nu \in M_\theta$ implies $\nu^S \in M_\theta$, we have

$$\phi\left(\tfrac{1}{2}\bar{\nu} + \tfrac{1}{2}\bar{\nu}^S\right) \le \tfrac{1}{2}\phi(\bar{\nu}) + \tfrac{1}{2}\phi(\bar{\nu}^S) = \phi(\bar{\nu}) \le \phi\left(\tfrac{1}{2}\bar{\nu} + \tfrac{1}{2}\bar{\nu}^S\right), \tag{3.29}$$

where the last inequality uses that $\bar{\nu}$ is a minimizer. Hence

$$\phi\left(\tfrac{1}{2}\bar{\nu} + \tfrac{1}{2}\bar{\nu}^S\right) = \tfrac{1}{2}\phi(\bar{\nu}) + \tfrac{1}{2}\phi(\bar{\nu}^S). \tag{3.30}$$

Via the form of ϕ in (1.7) this easily gives

$$\bar{\nu}(i,j) = \bar{\nu}^S(i,j) \quad \text{for all } i,j, \tag{3.31}$$

where we use that $\bar{\nu} \in M_\theta$ implies $\hat{\bar{\nu}} = \hat{\bar{\nu}}^S$ because of the condition $\sum_j \nu(i,j) = \sum_j \nu(j,i)$ for all i.

(ii) First deduce from the definition of ϕ that for a minimizer $\bar{\nu}$ and an admissible variation ∂ we have

$$\begin{aligned}
&\lim_{t\downarrow 0} \frac{1}{t}\left[\phi(\bar{\nu} + t\partial) - \phi(\bar{\nu})\right] \\
&\quad = -\sum_{i,j} \partial(i,j)\log A(i,j) + \sum_{i,j:\,\bar{\nu}(i,j)>0} \partial(i,j)\log\left(\frac{\bar{\nu}(i,j)}{\hat{\bar{\nu}}(i)}\right) \\
&\qquad + \sum_{i,j:\,\hat{\bar{\nu}}(i)=0} \partial(i,j)\log\left(\frac{\partial(i,j)}{\hat{\partial}(i)}\right) \\
&\qquad + \lim_{t\downarrow 0} \sum_{\substack{i,j:\,\bar{\nu}(i,j)=0\\ \hat{\bar{\nu}}(i)>0}} \partial(i,j)\left[\log\left(\frac{t\partial(i,j)}{\hat{\bar{\nu}}(i)}\right) - 1\right].
\end{aligned} \tag{3.32}$$

Comparison with (3.14) shows that

$$\partial(i,j) = 0 \quad \text{for all } i,j \text{ with } \bar{\nu}(i,j) = 0,\ \hat{\bar{\nu}}(i) > 0. \tag{3.33}$$

We shall use this property to exclude $\bar{\nu}(i,j) = 0$ by contradiction, as follows.

Pick $(k,l) \ne (1,1)$, $k \le l$, such that $\bar{\nu}(k,l) = \bar{\nu}(l,k) > 0$, which is possible by Lemma 5(ii) because we assume $\theta < 1$. Pick ∂ as

$$\begin{aligned}
\partial(k,l+1) &= \partial(l+1,k) = 1 + 1_{\{k=l+1\}}, \\
\partial(k,l) &= \partial(l,k) = -2\left(1 + 1_{\{k=l\}}\right), \\
\partial(k,l-1) &= \partial(l-1,k) = 1 + 1_{\{k=l-1\}}, \\
\partial(i,j) &= 0 \quad \text{otherwise}.
\end{aligned} \tag{3.34}$$

One easily checks that this ∂ is admissible, that is, fulfills (3.11)–(3.13). It follows from (3.33) that $\bar{\nu}(k,l+1) = \bar{\nu}(l+1,k) > 0$ and $\bar{\nu}(k,l-1) = \bar{\nu}(l-1,k) > 0$. Hence, by induction, $\bar{\nu}(i,j) > 0$ for all i,j.

(iii) Choose variations of the form

$$\begin{aligned}
&\partial(i,j) = \partial(j,i), \\
&\partial(i,j) \ne 0 \quad \text{if } (i,j) \text{ or } (j,i) \in \{(k,l),(m,n),(p,q)\}, \\
&\partial(i,j) = 0 \quad \text{otherwise}.
\end{aligned} \tag{3.35}$$

Here $k, l, m, n, p, q \in \mathbb{N}$ are arbitrary. The admissibility of ∂ requires

$$\partial(k,l) + \partial(m,n) + \partial(p,q) = 0, \tag{3.36}$$

$$(k+l)\partial(k,l) + (m+n)\partial(m,n) + (p+q)\partial(p,q) = 0. \tag{3.37}$$

[Note that because $\bar{\nu} > 0$ we no longer have the sign restriction in (3.11).] Now, combining (3.14) and (3.32) we have

$$\sum_{i,j} \partial(i,j) \log\left(\frac{\bar{\nu}(i,j)}{\hat{\bar{\nu}}(i) A(i,j)} \right) \geq 0, \tag{3.38}$$

which, in the notation of (3.15) and after using (3.35), gives

$$\bar{\eta}(k,l)\partial(k,l) + \bar{\eta}(m,n)\partial(m,n) + \bar{\eta}(p,q)\partial(p,q) \geq 0. \tag{3.39}$$

Next substitute (3.36) into (3.37) and (3.39) to eliminate $\partial(p,q)$. This yields the following implication, valid for all $\partial(k,l), \partial(m,n) \in \mathbb{R}$:

$$\begin{aligned} &(k+l-p-q)\partial(k,l) + (m+n-p-q)\partial(m,n) = 0, \\ \Rightarrow &(\bar{\eta}(k,l) - \bar{\eta}(p,q))\partial(k,l) + (\bar{\eta}(m,n) - \bar{\eta}(p,q))\partial(m,n) \geq 0. \end{aligned} \tag{3.40}$$

Hence there must exist $r \in \mathbb{R}$ such that

$$\frac{\bar{\eta}(k,l) - \bar{\eta}(p,q)}{k+l-p-q} = \frac{\bar{\eta}(m,n) - \bar{\eta}(p,q)}{m+n-p-q} = r. \tag{3.41}$$

Now pick $(p,q) = (1,1)$ to read off (3.16). □

3.4. *Variation over θ.*

PROPOSITION 5. *The minimizer $\bar{\theta}$ of* (1.5) *is the unique solution of*

$$\lambda(r(\bar{\theta})) = 1 \tag{3.42}$$

with $r(\theta)$ defined in (3.18). *The minimum is*

$$K(\bar{\theta}) = \frac{r(\bar{\theta})}{\bar{\theta}}. \tag{3.43}$$

PROOF. Differentiate (3.19) w.r.t. θ and use (3.18) to compute

$$\begin{aligned} \frac{d}{d\theta}(\theta K(\theta)) &= \log\frac{1}{\lambda(r)} + \frac{dr}{d\theta}\left[1 - \theta\frac{\lambda'(r)}{\lambda(r)}\right] = \log\frac{1}{\lambda(r)}, \\ \left(\frac{d}{d\theta}\right)^2(\theta K(\theta)) &= -\frac{\lambda'(r)}{\lambda(r)}\frac{dr}{d\theta} = -\frac{1}{\theta}\frac{dr}{d\theta}. \end{aligned} \tag{3.44}$$

This proves (3.42) [recall that $r = r(\theta)$ is decreasing in θ]. Because A_r is strictly increasing in r elementwise, it follows from the representation in (1.11) that $r \to \lambda(r)$ is strictly increasing with limits 0 as $r \downarrow -\infty$ and ∞ as $r \uparrow \infty$. Consequently, the solution of (3.42) exists and is unique. Substitution into (3.19) yields (3.43). □

3.5. *Proof of Theorems* 2 *A and B.*

PROOF OF THEOREM 2A. Combine Propositions 4 and 5. □

PROOF OF THEOREM 2B. Denote by $\lambda(\beta, r)$ the largest eigenvalue of the matrix $A_{\beta, r}$ belonging to the function βg, that is,

$$A_{\beta, r}(i, j) = e^{r(i+j-1)} e^{-\beta g(i+j-1)} P(i, j). \tag{3.45}$$

Denote by $r^*(\beta)$ and $\theta^*(\beta)$ the corresponding quantities defined in (1.13) and (1.14). By the same argument as in BCGH, Section 3.3, proving that $r \to \lambda(r)$ is analytic, we obtain here that $\beta \to \lambda(\beta, r)$ and $r \to \lambda(\beta, r)$ are analytic on $\{(\beta, r) \in \mathbb{C}^2\colon \operatorname{Re} \beta > 0\}$. Hence $(\beta, r) \to \lambda(\beta, r)$ is analytic [Hörmander (1966), Theorem 2.2.8]. It follows from the implicit function theorem [Hörmander (1966), Theorem 2.1.2] that $\beta \to r^*(\beta)$ and $\beta \to \theta^*(\beta)$ are analytic on $\{\beta \in C\colon \operatorname{Re} \beta > 0\}$.

To see that $\lim_{\beta \downarrow 0} \theta^*(\beta) = 0$, argue as follows. Let x_ξ be the vector $x_\xi(i) = \xi^{i-1}$ $(i \in \mathbb{N})$. From the identity

$$\sum_{j \in \mathbb{N}} \binom{i+j-2}{i-1} a^{j-1} = \frac{1}{(1-a)^i}, \qquad i \in \mathbb{N}, a \in [0, 1), \tag{3.46}$$

and from (1.3) one easily deduces that at $\beta = 0$,

$$A_{0, r} x_\xi = T(r, \xi) x_{T(r, \xi)} \tag{3.47}$$

with $T(r, \xi) = (2e^{-r} - \xi)^{-1}$. It follows that $\lambda(0, r)$ is the solution of the equation $\xi = T(r, \xi)$), that is,

$$\lambda(0, r) = e^{-r} - (e^{-2r} - 1)^{1/2}, \qquad r \leq 0. \tag{3.48}$$

Next pick $\varepsilon > 0$. Because $r \to \log \lambda(\beta, r)$ is convex for every $\beta > 0$ [see below (3.27)], we have [$\lambda_{(r)}$ denotes $\partial\lambda/\partial r$]

$$\frac{1}{\theta^*(\beta)} = \frac{\lambda_{(r)}}{\lambda}(\beta, r^*(\beta)) \geq \frac{\lambda_{(r)}}{\lambda}(\beta, -\varepsilon), \tag{3.49}$$

where we use that $r^*(\beta) > 0$ for every $\beta > 0$. As $\beta \downarrow 0$ the r.h.s. of (3.49) tends to $(\lambda_{(r)}/\lambda)(0, -\varepsilon)$ because $\beta \to \lambda(\beta, r)$ is analytic on $\{\beta \geq 0\}$ for every $r < 0$ (in analogy with BCGH, Lemma 11 and its proof). Now let $\varepsilon \downarrow 0$ and use that $\lambda(0, -\varepsilon) \to 1$, $\lambda_{(r)}(0, -\varepsilon) \to \infty$ by (3.48), to get the claim.

To see that $\lim_{\beta \to \infty} \theta^*(\beta) = 1$, argue as follows. Define

$$\hat{\lambda}(\beta, r) = \lambda(\beta, r + c\beta) \quad \text{with } c = g(1). \tag{3.50}$$

Because $g(i) > ig(1)$ for all $i > 1$, we have $A_{\beta, r + c\beta} \downarrow \frac{1}{2} e^r \delta_{(1,1)}$ pointwise as $\beta \uparrow \infty$ [note that $P_{(1,1)} = \frac{1}{2}$]. Hence,

$$\hat{\lambda}(\beta, r) \downarrow \tfrac{1}{2} e^r, \qquad \beta \uparrow \infty, \tag{3.51}$$

locally uniformly by Vitali's theorem. Define $\hat{r}^*(\beta)$ by

$$\hat{\lambda}(\beta, \hat{r}^*(\beta)) = 1. \tag{3.52}$$

Then (3.51) says that $\hat{r}^*(\beta) \to \log 2$ as $\beta \to \infty$. Because $\hat{r}^*(\beta) = r^*(\beta) - c\beta$, we have

$$(3.53) \qquad \frac{1}{\theta^*(\beta)} = \lambda_{(r)}(\beta, r^*(\beta)) = \hat{\lambda}_{(r)}(\beta, \hat{r}^*(\beta)).$$

As $\beta \to \infty$ the r.h.s. tends to $(\frac{1}{2}e^r)_{(r)}$ evaluated at $r = \log 2$, which equals 1. □

3.6. *Proof of Corollary* $*$. We know that

$$(3.54) \qquad 1 = \lambda(\beta, r^*(\beta)).$$

Therefore, using bracketed lower indices to denote partial derivatives, we have

$$(3.55) \qquad 0 = \lambda_{(\beta)}(\beta, r^*(\beta)) + r^*_{(\beta)}(\beta)\lambda_{(r)}(\beta, r^*(\beta)).$$

We also know that

$$(3.56) \qquad \frac{1}{\theta^*(\beta)} = \lambda_{(r)}(\beta, r^*(\beta)).$$

Hence,

$$(3.57) \qquad \left(\frac{1}{\theta^*(\beta)}\right)_{(\beta)} = \lambda_{(\beta, r)}(\beta, r^*(\beta)) + r^*_{(\beta)}(\beta)\lambda_{(r,r)}(\beta, r^*(\beta)).$$

Combining (3.55) and (3.57), and eliminating $r^*_{(\beta)}(\beta)$, we see that $\beta \to \theta^*(\beta)$ strictly increasing is equivalent to the inequality

$$(3.58) \qquad (\mathscr{L}\lambda)(\beta, r^*(\beta)) = \left[\lambda_{(r,r)}\lambda_{(\beta)} - \lambda_{(\beta,r)}\lambda_{(r)}\right](\beta, r^*(\beta)) > 0,$$

which is $(*)$.

We want now to make it plausible why (3.58) might be true. Recall (1.11), which says that

$$(3.59) \qquad \lambda(\beta, r) = \max_{\|x\|_2 = 1,\ x > 0} \langle x, A_{\beta,r}x\rangle.$$

We shall show that if $i \to g(i)/i$ is strictly increasing, then

$$(3.60) \qquad \mathscr{L}\langle x, A_{\beta,r}x\rangle > 0 \quad \text{for all } x > 0,\ x \in l^2(\mathbb{N}).$$

This is done as follows. The form of $A_{\beta,r}$ in (3.45) allows us to write partial derivatives w.r.t. β and r as moments. Indeed, dropping the indices β, r from A for notational convenience, we have

$$(3.61) \qquad \begin{aligned} \mathscr{L}\langle x, Ax\rangle &= \langle x, (A)_{(r,r)}x\rangle\langle x, (A)_{(\beta)}x\rangle - \langle x, (A)_{(\beta,r)}x\rangle\langle x, (A)_{(r)}x\rangle \\ &= \sum_{i,j,k,l} x(i)x(j)x(k)x(l)A(i,j)A(k,l) \\ &\quad \times\{-(i+j-1)^2 g(k+l-1) \\ &\qquad + (i+j-1)g(i+j-1)(k+l-1)\}. \end{aligned}$$

Symmetrize the term between braces using the abbreviation $a = i + j - 1$ and $b = k + l - 1$, that is,

$$\frac{1}{2}\{-a^2 g(b) - b^2 g(a) + ag(a)b + bg(b)a\}. \tag{3.62}$$

The latter may be rewritten as

$$\frac{1}{2}ab(a-b)\left(\frac{g(a)}{a} - \frac{g(b)}{b}\right). \tag{3.63}$$

Hence (3.63) is strictly positive whenever $a \neq b$. This proves that (3.61) is strictly positive (note that $A > 0$ and $x > 0$) and hence gives (3.60).

Although (3.60) supports (3.58), it is not enough to settle (3.58) because the class of functions $\{f(\beta, r)\colon \mathscr{L}f \geq 0\}$ is not closed under taking suprema (even without worrying about the differentiability). The operator $\mathscr{L}$ has the remarkable property that $\mathscr{L}\phi(f) = (\phi')^2\mathscr{L}f$ for every $\phi\colon \mathbb{R} \to \mathbb{R}$ differentiable. Therefore, using the fact that $\lambda(\beta, r) = \lim_{n\to\infty}(1/n)\log\langle x, A^n_{\beta,r}x\rangle$ for all $x > 0$, $x \in l^2(\mathbb{N})$, it follows that to get $\mathscr{L}\lambda \geq 0$ it would suffice to prove that $\mathscr{L}\langle x, A^n_{\beta,r}x\rangle \geq 0$ for all n. We have not been able to do this. □

4. Proof of Theorem 3. Theorem 3 is a consequence of the following stronger result. Recall the notation of (1.22)–(1.25).

PROPOSITION 6. *For every function g satisfying* (1.1), *every $\varepsilon > 0$, every function δ with $\delta(n) = o(n)$ and n sufficiently large*,

$$Q_n^g\left(\hat{\theta}_n \notin U^\varepsilon(\theta^*) | \hat{\theta}_n > 0\right) \leq \exp(-\delta(n)), \tag{4.1}$$

$$Q_n^g\left(\hat{\phi}_n \notin U^\varepsilon(\theta^*) | \hat{\theta}_n > 0\right) \leq \exp(-\delta(n)), \tag{4.2}$$

$$Q_n^g\left(\hat{\nu}_n^+ \notin U^\varepsilon(\nu^*) | \hat{\theta}_n > 0\right) \leq \exp(-\delta(n)), \tag{4.3}$$

where $U^\varepsilon(\theta^*) = (\theta^* - \varepsilon, \theta^* + \varepsilon)$ and $U^\varepsilon(\nu^*) = \{\nu\colon \|\nu - \nu^*\| < \varepsilon\}$ with $\|\cdot\|$ any metric inducing the weak topology.

The same relations as in (4.1)–(4.3) *hold conditioned on $\hat{\theta}_n < 0$, once we replace θ^* by $-\theta^*$ and $\hat{\nu}_n^+$ by $\hat{\nu}_n^-$.*

PROOF. For the proof of Proposition 6 it will be important that the minimizers $\bar{\theta} = \theta^*$ and $\bar{\nu} = \nu^*$ of (1.5) and (1.6) are attained and are unique, a fact which we have established in Section 3. In order to use these properties we shall first have to sharpen a bit the estimates in Section 2. [This is needed to translate the statements on the n-step path S^n into statements on the total local time process $\{l(x), x \in \mathbb{Z}\}$.]

LEMMA 6. *In Lemma* 1 (2.10) *and Lemma* 2 (2.17) *the symbol* $\asymp$ *may be replaced by* $= \{O(\exp(-\frac{1}{2}c\delta(n)))\} \times$, *with $c = g(2) - 2g(1) > 0$ and $\delta(n)$ in*

(2.1) *satisfying* $\delta(n)/(n\log n)^{1/2} \to \infty$ *as* $n \to \infty$, *provided the condition* $\hat{\theta}_n > 0$ *is added.*

PROOF. First note the obvious symmetry between a path and its reflection in 0. This allows us to condition on $\hat{\theta}_n > 0$. Next strengthen (2.13) to

$$G(T_n(S^n)) \le G(S^n) - c\delta(n) \quad \text{for all } S^n \in \Omega^n \setminus A_n^1. \tag{4.4}$$

This relation holds because each $S^n \in \Omega^n \setminus A_n^1$ gets reflected at least once under T_n, each reflection splits local times of at least $\delta(n)$ sites and $g(k+l) - g(k) - g(l) \ge g(2) - 2g(1)$ for all $k, l \ge 1$ by (1.1)(i) and (ii).

Next use (4.4) to strengthen (2.15) as follows:

$$\begin{aligned} &E^n\big(\exp[-G(S^n)]1\{S^n \in \Omega^n \setminus A_n^1\}\big) \\ &\quad\le \exp(-c\delta(n))E^n\big(\exp[-G(T_n(S^n))]1\{S^n \in \Omega^n \setminus A_n^1\}\big) \\ &\quad\le (2n)^{n/2\delta(n)}\exp(-c\delta(n))E^n\big(\exp[-G(S'^n)]1\{S'^n \in A_n^1\}\big). \end{aligned} \tag{4.5}$$

The first factor in the r.h.s. of (4.5) is bounded above by $\exp(-\frac{1}{2}c\delta(n))$ when $\delta(n)/(n\log n)^{1/2} \to \infty$ $(n \to \infty)$. This completes the proof of the claim concerning Lemma 1.

The claim concerning Lemma 2 is trivial because (2.18) shows that paths in $A_n^1 \setminus A_n^2$ have a contribution that decays faster than exponentially. □

Next we bring the large deviation analysis into play. Recall (1.24) and (1.25) as well as (1.5) and (1.6). The uniqueness of the minimizers $\bar{\theta}$ and $\bar{\nu}$ allows us to derive the next two lemmas.

LEMMA 7.

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n}\log E^n\Bigg(&\exp\Big[-\sum_{x\in\mathbb{Z}} g(l_n(x))\Big] \\ &\times 1\{\hat{\theta}_n \notin U^{\varepsilon}(\theta^*),\, S^n \in A_n^2\}\Bigg), \qquad (\hat{\theta}_n > 0) \\ = \inf_{\theta\in(0,1]\setminus U^{\varepsilon}(\theta^*)} &\theta K(\theta). \end{aligned} \tag{4.6}$$

PROOF. Equation (4.6) is exactly the kind of problem we treated in Section 2.3, the only difference being the additional indicator $1\{\hat{\theta}_n \notin U^{\varepsilon}(\theta^*)\}$. In order to modify the analysis of Section 2.3 accordingly, all that is needed is the observation that Proposition 1 carries over after we replace in (2.22) the domain of the infimum by $(0,1] \setminus U^{\varepsilon}(\theta^*)$. In particular, recall (2.25) where now the integral over θ runs over $(\delta(n)/n, 1] \setminus U^{\varepsilon}(\theta^*)$. □

Similarly as for the drift $\hat{\theta}_n$, we have for the empirical pair distribution $\hat{\nu}_n^+$ the following lemma.

LEMMA 8.

$$\limsup_{n\to\infty} \frac{1}{n} \log E^n\left(\exp\left[-\sum_{x\in\mathbb{Z}} g(l_n(x))\right] \times 1\{\hat{\nu}_n^+ \notin U^\varepsilon(\nu^*), S^n \in A_n^2\} | \hat{\theta}_n > 0\right) \le -\inf_{\theta\in(0,1]} \theta\left[\inf_{\nu\in M_\theta\setminus U^{\varepsilon/2}(\nu^*)} \phi(\nu)\right]. \tag{4.7}$$

PROOF. Do a similar argument as for Lemma 7. The key is the following modification of the upper bound in the proof of Lemma 3 in Section 2.3. On the set $A_n^3(\theta)$ we want to compare the two empirical distributions

$$\hat{\nu}_n^+ = \frac{1}{|R_n|} \sum_{x\in R_n} \delta_{(m_n^+(x-1), m_n^+(x))}, \tag{4.8}$$

$$\tilde{\nu}_n^+ = \frac{1}{|J_n(\theta)|} \sum_{x\in J_n(\theta)} \delta_{(m^+(x-1), m^+(x))}. \tag{4.9}$$

[Recall (1.22), (1.23) and (2.49).] Note that $I_n(\theta) \subseteq R_n \subseteq J_n(\theta)$ on the set $A_n^3(\theta)$. The modified estimate of (2.36) now reads as follows:

$$\begin{aligned} &E_\theta\left(\exp[-G(S^n)] 1\{\hat{\nu}_n^+ \notin U^\varepsilon(\nu^*), S \in A_n^3(\theta)\}\right) \\ &\quad \le E_\theta\left(\exp[-G(J_n(\theta), S)] \right. \\ &\qquad \left. \times 1\{\tilde{\nu}_n^+ \notin U^{\varepsilon/2}(\nu^*), S \in A(J_n(\theta), n + 2\delta(n))\}\right). \end{aligned} \tag{4.10}$$

This estimate uses the observation that for all $S \in A_n^3(\theta)$, in analogy with (2.34),

$$m_n^+(x, S^n) = \begin{cases} m^+(x, S), & \text{for } x \in I_n(\theta) \setminus \{\lfloor \theta n \rfloor\}, \\ m^+(x, S) - 1, & \text{for } x \in (J_n(\theta) \cup \{\lfloor \theta n \rfloor\}) \\ & \quad \setminus (I_n(\theta) \cup \{\lfloor \theta n \rfloor + \delta(n)\}), \\ 0, & \text{for } x \notin J_n(\theta) \setminus \{\lfloor \theta n \rfloor + \delta(n)\}. \end{cases} \tag{4.11}$$

The latter implies

$$\|\hat{\nu}_n^+ - \tilde{\nu}_n^+\|_{\text{var}} \le 2|J_n(\theta) \setminus R_n| / |J_n(\theta)| \le 4\delta(n)/(n + 2\delta(n)). \tag{4.12}$$

The topology induced by $\|\cdot\|_{\text{var}}$ is stronger than the topology induced by $\|\cdot\|$. Therefore, using the fact that $\delta(n) = o(n)$, we see that (4.12) implies for n sufficiently large

$$\hat{\nu}_n^+ \notin U^\varepsilon(\nu^*) \quad \Rightarrow \quad \tilde{\nu}_n^+ \notin U^{\varepsilon/2}(\nu^*). \tag{4.13}$$

This explains (4.10). The rest of the argument is the same as for (4.6) and we omit further details. □

We are now ready to prove Proposition 6. We start by proving (4.1). Write

$$\begin{aligned}(4.14)\qquad & E^n\Big(\exp[-G(S^n)]1\{\hat{\theta}_n \notin U^\varepsilon(\theta^*)\}|\hat{\theta}_n > 0\Big)\\ &\quad \le E^n\Big(\exp[-G(S^n)]1\{\hat{\theta}_n \notin U^\varepsilon(\theta^*), S^n \in A_n^2\}|\hat{\theta}_n > 0\Big)\\ &\qquad + E^n\Big(\exp[-G(S^n)]1\{S^n \in \Omega^n \setminus A_n^2\}|\hat{\theta}_n > 0\Big).\end{aligned}$$

By Lemma 7 the first term in the r.h.s. of (4.14) is bounded above for large n by

$$(4.15)\qquad \exp(-(z + \varsigma(\varepsilon))n) \quad \text{for some } \varsigma(\varepsilon) > 0,$$

because $\theta^* = \bar{\theta}$ is the unique minimizer of (1.5). By Lemma 6 the second term in the r.h.s. of (4.14) is for large n bounded above by

$$(4.16)\qquad O\big(\exp(-\tfrac{1}{2}c\delta(n))\big)E^n(\exp[-G(S^n)]) = O\big(\exp(-\tfrac{1}{2}c\delta(n))\big)Z_n^g.$$

It now follows from the definition of Q_n^g [recall (0.1) and (0.2)], and from (1.4) in Theorem 1 combined with (4.14)–(4.16), that

$$(4.17)\qquad Q_n^g\big(\hat{\theta}_n \notin U^\varepsilon(\theta^*)|\hat{\theta}_n > 0\big) \le \exp\big(-\tfrac{1}{2}\varsigma(\varepsilon)n\big) + O\big(\exp\big(-\tfrac{1}{2}c\delta(n)\big)\big).$$

This proves the assertion in (4.1), because $\delta(n)$ can be picked arbitrarily close to $o(n)$.

The proof of (4.2) uses (4.1) and the fact that the main contribution to the expectation defining Z_n^g comes from $\{S^n \in A_n^2\}$ (recall Lemma 6). On this event, we have $|\hat{\theta}_n - \hat{\phi}_n| < 2\delta(n)/n \to 0$ $(n \to \infty)$.

The proof of (4.3) proceeds via the same arguments used to prove (4.1). For every $\theta \in (0, 1]$ the minimizer $\bar{\nu}(\theta)$ of (1.6) is unique, so that the r.h.s. of (4.7) can be bounded above by $-(z + \varsigma'(\varepsilon))$ for some $\varsigma'(\varepsilon) > 0$. To see the latter, proceed by contradiction, using the fact that ϕ is lower semicontinuous and the infimum can be restricted to a compact subset of M_θ (recall Lemma 4). □

Acknowledgment. The research for this paper was carried out while the second author was a guest at the Institut für Mathematische Stochastik of the Universität Göttingen in the spring and summer of 1991.

REFERENCES

BAILLON, J.-B., CLÉMENT, PH., GREVEN, A. and DEN HOLLANDER, F. (1991). On a variational problem for an infinite particle system in a random medium. Part I: The global growth rate. Preprint.

BAILLON, J.-B., CLÉMENT, PH., GREVEN, A. and DEN HOLLANDER, F. (1993). A variational approach to branching random walk in random environment. *Ann. Probab.* **21** 290–317.

BOLTHAUSEN, E. (1990). On self-repellent one-dimensional random walks. *Probab. Theory Related Fields* **86** 423–441.

BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, MA.

CZECH, R. and KEHR, K. W. (1984). Spin depolarization for one-dimensional random walk. *Phys. Rev. Lett.* **53** 1783–1786.

CZECH, R. and KEHR, K. W. (1986). Depolarization of rotating spins by random walks on lattices. *Phys. Rev. B* **34** 261–277.

DEUSCHEL, J.-D., and STROOCK, D. W. (1989). *Large Deviations*. Academic, Boston.

DONSKER, M. D. and VARADHAN, S. R. S. (1975). Asymptotics for the Wiener sausage. *Commun. Pure Appl. Math.* **28** 525–565.

EISELE, TH. and LANGE, R. (1987). Asymptotics for the Wiener sausage with drift. *Probab. Theory Related Fields* **74** 125–140.

ELLIS, R. (1984). *Entropy, Large Deviations and Statistical Mechanics. Grundlehren der Mathematischen Wissenschaften* **271**. Springer, New York.

GREVEN, A. and DEN HOLLANDER, F. (1992). Branching random walk in random environment: Phase transitions for local and global growth rates. *Probab. Theory Related Fields* **91** 195–249.

HÖRMANDER, L. (1966). *An Introduction to Complex Analysis in Several Variables*. Van Nostrand, Princeton, NJ.

KUSUOKA, S. (1985). Asymptotics of the polymer measure in one dimension. In *Infinite Dimensional Analysis and Stochastic Processes* (S. Albeverio, ed.) *Research Notes Math.* **124** 66–82. Pitman, Boston.

MAZO, R. M. and VAN DEN BROECK, C. (1986). Spectrum of an oscillator making a random walk in frequency. *Phys. Rev. A* **34** 2364–2374.

WESTWATER, J. (1984). On Edwards' model for polymer chains. In *Trends and Developments in the Eighties* (S. Albeverio and P. Blanchard, eds.). *Bielefeld Encounters in Math. Phys.* **4 / 5**. World Science, Singapore.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT GÖTTINGEN
LOTZESTRASSE 13
D-37083 GÖTTINGEN
GERMANY

MATHEMATISCH INSTITUUT
UNIVERSITEIT UTRECHT
P.O. BOX 80.010
NL-3508TA UTRECHT
THE NETHERLANDS