JOINT DENSITY FOR THE LOCAL TIMES OF CONTINUOUS-TIME MARKOV CHAINS: EXTENDED VERSION

David Brydges[1], Remco van der Hofstad[2] and Wolfgang König[3]

September 4, 2006

Abstract. We investigate the local times of a continuous-time Markov chain on an arbitrary discrete state space. For fixed finite range of the Markov chain, we derive an explicit formula for the joint density of all local times on the range, at any fixed time. We use standard tools from the theory of stochastic processes and finite-dimensional complex calculus.

We apply this formula in the following directions: (1) we derive large deviation upper estimates for the normalized local times beyond the exponential scale, (2) we derive the upper bound in Varadhan's lemma for any measurable functional of the local times, and (3) we derive large deviation upper bounds for continuous-time simple random walk on large subboxes of $\mathbb{Z}^d$ tending to $\mathbb{Z}^d$ as time diverges. We finally discuss the relation of our density formula to the Ray-Knight theorem for continuous-time simple random walk on $\mathbb{Z}$, which is analogous to the well-known Ray-Knight description of Brownian local times. In this extended version, we prove that the Ray-Knight theorem follows from our density formula.

1. Introduction

Let Λ be a finite or countably infinite set and let $A = (A_{x,y})_{x,y\in\Lambda}$ be the generator, sometimes called the Q-matrix, of a continuous-time Markov chain $(X_t)_{t\in[0,\infty)}$ on Λ. Under the measure $\mathbb{P}_a$, the chain starts at $X_0 = a \in \Lambda$, and by $\mathbb{E}_a$ we denote the corresponding expectation. The main object of our study are the *local times*, defined by

$$\ell_T(x) = \int_0^T \mathbb{1}_{\{X_s = x\}}\,\mathrm{d}s, \qquad x \in \Lambda, T > 0, \tag{1.1}$$

which register the amount of time the chain spends in x up to time T. We have $\langle \ell_T, V\rangle = \int_0^T V(X_s)\,\mathrm{d}s$ for any bounded function $V\colon \Lambda \to \mathbb{R}$, where $\langle\cdot,\cdot\rangle$ denotes the standard inner product on $\mathbb{R}^\Lambda$. We conceive the normalized local times tuple, $\frac{1}{T}\ell_T = (\frac{1}{T}\ell_T(x))_{x\in\Lambda}$, as a random element of the set $\mathcal{M}_1(\Lambda)$ of probability measures on Λ.

The local times tuple $\ell_T = (\ell_T(x))_{x\in\Lambda}$, and in particular its large-T behavior, are of fundamental interest in many branches and applications of probability theory. We are particularly interested in the large deviation of $\frac{1}{T}\ell_T$. A by now classical result [Gär77, DV75-83] states, for a *finite* state space Λ, a large deviation principle for $\frac{1}{T}\ell_T$, for any starting point $a \in \Lambda$, on the scale T. More precisely, for any closed set $\Gamma \subseteq \mathcal{M}_1(\Lambda)$,

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{P}_a\big(\tfrac{1}{T}\ell_T \in \Gamma\big) \leq -\inf_{\mu\in\Gamma} I_A(\mu), \tag{1.2}$$

[1]Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada. db5d@math.ubc.ca
[2]Department of Mathematics and Computer Science, Technical University Eindhoven, Post Box 513, 5600 MB Eindhoven, The Netherlands. rhofstad@win.tue.nl
[3]Mathematisches Institut, Universität Leipzig, Augustusplatz 10/11, D-04157 Leipzig, Germany. koenig@math.uni-leipzig.de

and, for any open set $G \subseteq \mathcal{M}_1(\Lambda)$,

$$\liminf_{T\to\infty} \frac{1}{T} \log \mathbb{P}_a\big(\tfrac{1}{T}\ell_T \in G\big) \geq -\inf_{\mu\in G} I_A(\mu). \tag{1.3}$$

The rate function I_A may be written

$$I_A(\mu) = -\inf\Big\{\Big\langle Ag, \frac{\mu}{g}\Big\rangle \,\Big|\, g\colon \Lambda \to (0,\infty)\Big\}. \tag{1.4}$$

In case that A is a symmetric matrix, $I_A(\mu) = \|(-A)^{\frac12}\sqrt{\mu}\|_2^2$ is equal to the Dirichlet form of A applied to $\sqrt{\mu}$. The topology used on $\mathcal{M}_1(\Lambda)$ is the weak topology induced by convergence of integrals against all bounded functions $\Lambda \to \mathbb{R}$, i.e., the standard topology of pointwise convergence since Λ is assumed finite. For infinite Λ, versions of this large deviations principle may be formulated for the restriction of the chain to some finite subset of Λ. A standard way of proving the above principle of large deviations is via the Gärtner-Ellis theorem; see [DZ98] for more background on large deviation theory. One of the major corollaries is Varadhan's lemma, which states that

$$\lim_{T\to\infty} \frac{1}{T} \log \mathbb{E}_x\big[\mathrm{e}^{TF(\frac{1}{T}\ell_T)}\big] = -\inf_{\mu\in\mathcal{M}_1(\Lambda)} \big[I_A(\mu) - F(\mu)\big], \tag{1.5}$$

for any function $F\colon \mathcal{M}_1(\Lambda) \to \mathbb{R}$ that is bounded and continuous in the above topology. We would like to stress that in many situations it is the upper bound in (1.5) that is difficult to prove since F often fails to be upper semicontinuous. (However, often F turns out to be lower semicontinuous or well approximated by lower semicontinous functions, so that the proof of the lower bound in (1.5) is often simpler.)

In the present paper, we considerably strengthen the above large deviation principle and the assertion in (1.5) by presenting an *explicit density of the random variable* ℓ_T, i.e., a joint density of the tuple $(\ell_T(x))_{x\in\Lambda}$, for any fixed $T > 0$. We do this for either a finite state space Λ or for the restriction to a finite subset. This formula opens up several new possibilities, such as

1. more precise asymptotics for the probabilities in (1.2) and (1.3) and for the expectation on the left of (1.5),
2. the validity of (1.5) for many discontinuous functions F,
3. versions of the large deviation principle for rescaled versions of the local times on state spaces $\Lambda = \Lambda_T$ coupled with T and growing to some infinite set,

Clearly, a closed analytical formula for the density of the local times is quite interesting in its own right. Unfortunately, our expression for the local times density is rather involved and is quite hard to evaluate asymptotically. Actually, not even the nonnegativity of the density can be easily seen from our formula. Luckily, *upper bounds* on the density are more easily obtained. We will be able to use these upper bounds to derive proofs of (1.2) and of the upper bound in (1.5) for *every measurable set* Γ, respectively, for *every measurable function* F, which is a great improvement.

This paper is organized as follows. In Section 2, we identify the density of the local times in Theorem 2.1, and prove Theorem 2.1. In Section 3, we use Theorem 2.1 to prove large deviation upper bounds in Theorem 3.6. Finally, we close in Section 4 by discussing our results, by relating them to the history of the problem and by discussing the relation to the Ray-Knight theorem.

2. Density of the local times

In this section, we present our fundamental result, Theorem 2.1, which is the basis for everything that follows. By

$$R_T = \operatorname{supp}(\ell_T) = \{X_s\colon s \in [0,T]\} \subseteq \Lambda \tag{2.1}$$

we denote the *range* of the Markov chain. Note that given $\{R_T \subseteq R\}$ for some finite set $R \subseteq \Lambda$, the random tuple $(\ell_T(x))_{x\in R}$ does not have a density with respect to the Lebesgue measure, since the event $\{\ell_T(x)=0\}$ occurs with positive probability for any $x\in R$, except for the initial site of the chain. However, given $\{R_T=R\}$ for some $R\subseteq\Lambda$, the tuple $(\ell_T(x))_{x\in R}$ takes values in the simplex

$$\mathcal{M}_T^+(R)=\Big\{l\colon R\to(0,\infty)\,\Big|\,\sum_{x\in R}l(x)=T\Big\},\tag{2.2}$$

which is a convex open subset of the hyperplane in $\mathbb{R}^R$ that is perpendicular to $\mathbb{1}$. It will turn out that on $\{R_T=R\}$, the tuple $(\ell_T(x))_{x\in R}$ has a density with respect to the Lebesgue measure σ_T on $\mathcal{M}_T^+(R)$ defined by the disintegration of Lebesgue measure into surface measures,

$$\int \mathrm{d}^R l\, F(l)=\int_0^\infty \mathrm{d}T\,\int_{\mathcal{M}_T^+}\sigma_T(\mathrm{d}l)\;F(l),\tag{2.3}$$

where $F\colon(0,\infty)^R\to\mathbb{R}$ is bounded and continuous with compact support.

We need some notation. Let $R\subseteq\Lambda$ and let $a,b\in R$. For a matrix $M=(M_{x,y})_{x,y\in\Lambda}$ we denote by $\det_{ab}^{(R)}(M)$ the (b,a) cofactor of the $R\times R$-submatrix of M, namely the determinant of the matrix $(1_{x\neq b}M_{x,y}1_{y\neq a}+1_{x=b,y=a})_{x,y\in R}$. We write $\det_{ab}$ instead of $\det_{ab}^{(\Lambda)}$ when no confusion can arise. By ∂_l we denote the $\Lambda\times\Lambda$-diagonal matrix with (x,x)-entry ∂_{l_x}, which is the partial derivative with respect to l_x. Hence, $\det_{ab}^{(R)}(M+\partial_l)$ is a linear differential operator of order $|R|-2+\delta_{a,b}$.

Then our main result reads as follows:

Theorem 2.1 (Density of the local times)**.** *Let Λ be a finite or countably infinite set with at least two elements and let $A=(A_{x,y})_{x,y\in\Lambda}$ be the conservative generator of a continuous-time Markov chain on Λ. Fix a finite subset R of Λ and sites $a,b\in R$. Then, for every $T>0$ and for every bounded measurable function $F\colon\mathcal{M}_T^+(R)\to\mathbb{R}$,*

$$\mathbb{E}_a\big[F(\ell_T)\mathbb{1}_{\{X_T=b\}}\mathbb{1}_{\{R_T=R\}}\big]=\int_{\mathcal{M}_T^+(R)}F(l)\rho_{ab}^{(R)}(l)\,\sigma_T(\mathrm{d}l),\tag{2.4}$$

where, for $l\in\mathcal{M}_T^+(R)$,

$$\rho_{ab}^{(R)}(l)=\det\nolimits_{ab}^{(R)}\big(-A+\partial_l\big)\int_{[0,2\pi]^R}\mathrm{e}^{\sum_{x,y\in R}A_{x,y}\sqrt{l_x}\sqrt{l_y}\mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}}\prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}.\tag{2.5}$$

Alternative expressions for the density $\rho_{ab}^{(R)}$ are found in Proposition 2.5 below. Note that the density $\rho_{ab}^{(R)}$ does not depend on the values of the generator outside R, nor on T. The formula for the density is explicit, but quite involved, in particular as it involves determinants of large matrices, additional multiple integrals, and various partial derivatives. For example, it is not clear from (2.5) that $\rho_{ab}^{(R)}$ is non-negative. Nevertheless, the formula allows us to prove rather precise and transparent large deviation upper bounds for the local times as we shall see later. As we will discuss in more detail in Section 4, Theorem 2.1 finds its roots in the work of Luttinger [Lut83] who expressed expectations of functions of the local times in terms of integrals in which there are "functions" of anticommuting differential forms (Grassman variables). It is not clear from his work that the Grassman variables can be removed without creating intractable expressions. Theorem 2.1 accomplishes this removal. We also provide a proof that makes no overt use of Grassman variables; the determinant is their legacy.

To prepare for the proof, we need the following two lemmas and some notation. We write $\phi=u+\mathrm{i}v$ and $\overline{\phi}=u-\mathrm{i}v$, where $u,v\in\mathbb{R}^\Lambda$, and we use $\mathrm{d}^\Lambda u\,\mathrm{d}^\Lambda v$ to denote the Lebesgue measure on $\mathbb{R}^\Lambda\times\mathbb{R}^\Lambda$. Let $\langle\phi,\psi\rangle=\sum_{x\in\Lambda}\phi_x\psi_x$ be the real inner product on $\mathbb{C}^\Lambda$.

Lemma 2.2. *Let Λ be a finite set, and let $M\in\mathbb{C}^{\Lambda\times\Lambda}$. If $\Re\langle\phi,M\overline{\phi}\rangle>0$ for any $\phi\in\mathbb{C}^\Lambda\setminus\{0\}$, then*

$$\int_{\mathbb{R}^\Lambda\times\mathbb{R}^\Lambda}\mathrm{d}^\Lambda u\,\mathrm{d}^\Lambda v\,\mathrm{e}^{-\langle\phi,M\overline{\phi}\rangle}=\frac{\pi^{|\Lambda|}}{\det(M)}.\tag{2.6}$$

Remark 2.3. By introducing polar coordinates $(l, \theta) \in [0,\infty)^\Lambda \times [0, 2\pi]^\Lambda$ via

$$\phi_x = \sqrt{l_x} \mathrm{e}^{\mathrm{i}\theta_x}, \qquad x \in \Lambda, \tag{2.7}$$

we can transform

$$\mathrm{d}^\Lambda u \, \mathrm{d}^\Lambda v = \pi^{|\Lambda|} \prod_{x\in\Lambda} \Big(\mathrm{d}l_x \, \frac{\mathrm{d}\theta_x}{2\pi}\Big) = 2^{-|\Lambda|} \mathrm{d}^\Lambda l \, \mathrm{d}^\Lambda \theta \tag{2.8}$$

and can rewrite (2.6) in the form

$$\int_{[0,\infty)^\Lambda \times [0,2\pi]^\Lambda} \mathrm{d}^\Lambda l \, \frac{\mathrm{d}^\Lambda \theta}{(2\pi)^{|\Lambda|}} \, \mathrm{e}^{-\langle \phi, M\overline{\phi}\rangle} = \frac{1}{\det(M)}. \tag{2.9}$$

$\Diamond$

Proof of Lemma 2.2. We define the complex inner product $(\phi,\psi) = \langle \phi, \overline{\psi}\rangle$. Any unitary matrix $U \in \mathbb{C}^{\Lambda\times\Lambda}$ defines a complex linear transformation on $\mathbb{C}^\Lambda$ by $\phi' = U\phi$. By writing $\phi = u + \mathrm{i}v$ and $\phi' = u' + \mathrm{i}v'$ we obtain a real linear transformation $\widetilde{U} \colon (u, v) \mapsto (u', v')$ on $\mathbb{R}^\Lambda \oplus \mathbb{R}^\Lambda$. The map $\widetilde{U}$ is orthogonal, because

$$\langle u', u'\rangle + \langle v', v'\rangle = (\phi', \phi') = (U\phi, U\phi) = (\phi, \phi) = \langle u, u\rangle + \langle v, v\rangle.$$

Let M^* be the adjoint to M so that $(\phi, M\psi) = (M^*\phi, \psi)$. First we consider the case where $M = M^*$. The hypothesis $\Re\langle \phi, M\overline{\phi}\rangle > 0$ can be rewritten as $(\phi, M\phi) > 0$, so that M has throughout positive eigenvalues λ_x, $x \in \Lambda$. Since M is self-adjoint there exists a unitary transformation U such that $U^*MU = D$, where D is diagonal with diagonal entries $D_{x,x} = \lambda_x > 0$. Thus, by the change of variables $(u', v') = \widetilde{U}(u, v)$,

$$\int \mathrm{d}^\Lambda u \, \mathrm{d}^\Lambda v \, \mathrm{e}^{-(\phi, M\phi)} = \int \mathrm{d}^\Lambda u \, \mathrm{d}^\Lambda v \, \mathrm{e}^{-(\phi, D\phi)}.$$

The integral on the right hand side factors into a product of integrals

$$\prod_{x\in\Lambda} \int_{\mathbb{R}} \mathrm{d}u \int_{\mathbb{R}} \mathrm{d}v \, \mathrm{e}^{-\lambda_x u^2 - \lambda_x v^2} = \prod_{x\in\Lambda} \frac{\pi}{\lambda_x} = \frac{\pi^{|\Lambda|}}{\det(M)}.$$

The lemma is proved for the case $M = M^*$.

Now we turn to the case where $M^* \neq M$. Let

$$S = \frac{1}{2}\big(M + M^*\big) \qquad \text{and} \qquad A = \frac{1}{2\mathrm{i}}\big(M - M^*\big).$$

Thus, S and A are self-adjoint and $M = S + \mathrm{i}A$. Also, $(\phi, S\phi) = \Re\langle \phi, M\overline{\phi}\rangle$ which is positive by the hypothesis. Therefore the eigenvalues of S are strictly positive.

For $\mu \in \mathbb{C}$ we define $M(\mu) = S+\mu A$. For μ real, the matrix $M(\mu)$ is self-adjoint. Observe that $M(\mu)$ has throughout strictly positive eigenvalues when $\mu = 0$. Hence, the real part of the characteristic polynomial of $M(\mu)$ is non-zero on $(-\infty, 0]$, and therefore bounded away from zero on $(-\infty, 0]$, for $\mu = 0$. By continuity of the real part of this polynomial in μ, the latter property persists to all μ in a suitable open interval $I \subset \mathbb{R}$ containing the origin. Therefore, $M(\mu)$ has throughout strictly positive eigenvalues for all $\mu \in I$. Thus we have $(\phi, M(\mu)\phi) > 0$ for all nonzero ϕ and all $\mu \in I$.

Now we apply the preceding with $M = M(\mu)$, and obtain, for $\mu \in I$,

$$\det(M(\mu)) \int \mathrm{d}^\Lambda u \, \mathrm{d}^\Lambda v \, \mathrm{e}^{-(\phi, M(\mu)\phi)} = \pi^{|\Lambda|}. \tag{2.10}$$

Both sides of this equation are analytic in μ for $\Re\mu \in I$ because $\det(M(\mu))$ is a polynomial in μ, and the integral of the analytic function $\exp(-(\phi, M(\mu)\phi))$ is analytic by Morera's theorem and the Fubini theorem, as well as the remark that

$$\big|\mathrm{e}^{-(\phi, M(\mu)\phi)}\big| = \big|\mathrm{e}^{-(\phi, S\phi) - \mu(\phi, A\phi)}\big| = \mathrm{e}^{-(\phi, S\phi) - \Re\mu(\phi, A\phi)} = \mathrm{e}^{-(\phi, M(\Re\mu)\phi)}.$$

By analytic continuation (2.10) holds for $\Re\mu \in I$ and in particular for $\mu = \mathrm{i}$. At $\mu = \mathrm{i}$, $M(\mu) = M$. $\square$

Lemma 2.4. *Let Λ be a finite set, let $M \in \mathbb{C}^{\Lambda\times\Lambda}$, and $v = (v_x)_{x\in\Lambda} \in \mathbb{C}^\Lambda$. Then, for any continuously differentiable function $g\colon \mathbb{C}^\Lambda \to \mathbb{R}$,*

$$\det_{ab}(M + \partial_l)\big(\mathrm{e}^{\langle v,\cdot\rangle} g\big)(l) = \mathrm{e}^{\langle v,l\rangle} \det_{ab}(M + V + \partial_l) g(l), \qquad l \in \mathbb{R}^\Lambda, \tag{2.11}$$

where $V = (\delta_{xy} v_x)_{x,y\in\Lambda}$ denotes the diagonal matrix with diagonal entries v_x.

Proof. By a cofactor expansion, one sees that, for any diagonal matrix W, $\det_{ab}(M + W) = \sum_{X\subseteq\Lambda\setminus\{a,b\}} c_X \prod_{x\in X} W_{x,x}$ for suitable coefficients c_X depending only on the entries of M. Analogously, $\det_{ab}(M + \partial_l) = \sum_{X\subseteq\Lambda\setminus\{a,b\}} c_X \partial_l^X$, where we used the notation $\partial_l^X = \prod_{x\in X} \partial_{l_x}$. Therefore,

$$\begin{aligned}\mathrm{e}^{-\langle v,l\rangle} \det_{ab}(M + \partial_l)\big(\mathrm{e}^{\langle v,\cdot\rangle} g\big)(l) &= \sum_{X\subseteq\Lambda\setminus\{a,b\}} c_X \mathrm{e}^{-\langle v,l\rangle} \partial_l^X \big(\mathrm{e}^{\langle v,\cdot\rangle} g\big)(l) \\ &= \sum_{X\subseteq\Lambda\setminus\{a,b\}} c_X \prod_{x\in\Lambda} \big(v_x + \partial_{l_x}\big) g(l) = \det_{ab}(M + V + \partial_l) g(l).\end{aligned}$$

$\square$

Proof of Theorem 2.1. We have divided the proof into six steps. In the first five steps we assume that Λ is a finite set, and we put $R = \Lambda$. Recall the notation in Remark 2.3, which will be used throughout this proof. We abbreviate $D_l = \det_{ab}(-A - \partial_l)$.

Step 1: For any $v \in \mathbb{C}^\Lambda$ with $\Re v \in (-\infty, 0)^\Lambda$, for $F(l) = \mathrm{e}^{\langle v,l\rangle}$,

$$\int_0^\infty \mathbb{E}_a\big[F(\ell_T) \mathbb{1}_{\{X_T=b\}}\big]\,\mathrm{d}T = \int_{[0,\infty)^\Lambda\times[0,2\pi]^\Lambda} (D_l F)(l) \mathrm{e}^{\langle\phi, A\overline{\phi}\rangle} \prod_{x\in\Lambda} \Big(\mathrm{d}l_x \frac{\mathrm{d}\theta_x}{2\pi}\Big). \tag{2.12}$$

Proof. Recall that $\langle v, \ell_T\rangle = \int_0^T v(X_s)\,\mathrm{d}s$ to obtain

$$\begin{aligned}\int_0^\infty \mathbb{E}_a\big[F(\ell_T) \mathbb{1}_{\{X_T=b\}}\big]\,\mathrm{d}T &= \int_0^\infty \mathbb{E}_a\Big[\mathrm{e}^{\int_0^T v(X_s)\,\mathrm{d}s} \mathbb{1}_{\{X_T=b\}}\Big]\,\mathrm{d}T = \int_0^\infty \big(\mathrm{e}^{T(A+V)}\big)_{a,b}\,\mathrm{d}T \\ &= (-A - V)^{-1}_{a,b},\end{aligned} \tag{2.13}$$

where V is the diagonal matrix with (x, x)-entry v_x, and $M_{x,y}$ denotes the (x, y)-entry of a matrix M. In order to see the last identity in (2.13), we note that

$$\int_0^\infty \big(\mathrm{e}^{T(A+V)}\big)_{a,b}\,\mathrm{d}T = \Big(\int_0^\infty \mathrm{e}^{T(A+V)}\,\mathrm{d}T\Big)_{a,b}, \tag{2.14}$$

and that

$$(A + V) \int_0^\infty \mathrm{e}^{T(A+V)}\,\mathrm{d}T = \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}T} \mathrm{e}^{T(A+V)}\,\mathrm{d}T = -I. \tag{2.15}$$

By Cramer's rule followed by (2.9),

$$\begin{aligned}(-A - V)^{-1}_{a,b} = \frac{\det_{ab}(-A - V)}{\det(-A - V)} &= \int \det_{ab}(-A - V) \mathrm{e}^{\langle\phi,(A+V)\overline{\phi}\rangle} \prod_{x\in\Lambda} \Big(\mathrm{d}l_x \frac{\mathrm{d}\theta_x}{2\pi}\Big) \\ &= \int \det_{ab}(-A - V) \mathrm{e}^{\langle v,l\rangle} \mathrm{e}^{\langle\phi,A\overline{\phi}\rangle} \prod_{x\in\Lambda} \Big(\mathrm{d}l_x \frac{\mathrm{d}\theta_x}{2\pi}\Big).\end{aligned} \tag{2.16}$$

We use Lemma 2.4 with $g = 1$ and $M = A$ to obtain that

$$\begin{aligned}\det_{ab}(-A - V)\mathrm{e}^{\langle v,l\rangle} &= (-1)^{|\Lambda|-1} \mathrm{e}^{\langle v,l\rangle} \det_{ab}(A + V) = (-1)^{|\Lambda|-1} \det_{ab}(A + \partial_l) \mathrm{e}^{\langle v,l\rangle} \\ &= \det_{ab}(-A - \partial_l) \mathrm{e}^{\langle v,l\rangle} = (D_l F)(l),\end{aligned} \tag{2.17}$$

where we recall that $D_l = \det_{ab}(-A - \partial_l)$. Substituting this in (2.16) and combining this with (2.13), we conclude that (2.12) holds. $\square$

Step 2: The formula (2.12) is also valid for functions F of the form

$$F(l) = \prod_{x\in\Lambda} \big(\mathrm{e}^{v_x l_x} f_x(l_x)\big), \qquad f_x \in \mathcal{C}^2\big((0,\infty)\big), \mathrm{supp}(f_x) \subseteq (0,\infty) \text{ compact}, \Re v_x < 0. \tag{2.18}$$

Proof. Note that (2.12) is linear in F and so if we know it for exponentials, then we obtain it for linear combinations of exponentials. In more detail, consider the Fourier representation $f_x(l_x) = \int_{\mathbb{R}} \widehat{f}_x(w_x)\mathrm{e}^{\mathrm{i}w_x l_x}\,\mathrm{d}w_x$. Apply (2.12) for v replaced by $v + \mathrm{i}w$ with $w \in \mathbb{R}^\Lambda$ to obtain

$$\int_0^\infty \mathbb{E}_a\big[\mathrm{e}^{\langle v,\ell_T\rangle}\mathrm{e}^{\mathrm{i}\langle w,\ell_T\rangle}1\!\!1_{\{X_T=b\}}\big]\,\mathrm{d}T = \int_{[0,\infty)^\Lambda\times[0,2\pi]^\Lambda} \Big(D_l \mathrm{e}^{\langle v+\mathrm{i}w,l\rangle}\Big)\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle} \prod_{x\in\Lambda}\Big(\mathrm{d}l_x\,\frac{\mathrm{d}\theta_x}{2\pi}\Big).$$

Now multiply both sides with $\prod_{x\in\Lambda}\widehat{f}_x(w_x)$ and integrate over $\mathbb{R}^\Lambda$ with respect to $\mathrm{d}^\Lambda w$. Then we apply Fubini's theorem to move the $\mathrm{d}^\Lambda w$ integration inside. From the representation

$$\widehat{f}_x(w_x) = \frac{1}{2\pi}\int_{\mathbb{R}} f_x(l_x)\mathrm{e}^{-\mathrm{i}w_x l_x}\,\mathrm{d}l_x$$

we see that $\widehat{f}_x$ is continuous by the dominated convergence theorem. Furthermore, $\widehat{f}_x$ satisfies the bound

$$|\widehat{f}_x(w_x)| = \Big|\frac{1}{2\pi(\mathrm{i}w_x)^2}\int f_x(l)\frac{\mathrm{d}^2}{\mathrm{d}w_x^2}\mathrm{e}^{-\mathrm{i}w_x l}\,\mathrm{d}l\Big| = \frac{1}{2\pi}\frac{1}{w_x^2}\Big|\int f_x''(l)\mathrm{e}^{-\mathrm{i}w_x l}\,\mathrm{d}l\Big| \le \frac{1}{2\pi}\frac{1}{w_x^2}\int|f_x''(l)|\,\mathrm{d}l.$$

Hence, all functions $w_x \mapsto \widehat{f}_x(w_x)$ are absolutely integrable, and the exponentials with $\Re v_x < 0$ make the integration over l_x convergent for any $x \in R$. $\square$

In the following we abbreviate $D_l^* = \det_{ab}(-A + \partial_l)$.

Step 3: For F as in (2.18),

$$\int_0^\infty \mathbb{E}_a\big[F(\ell_T)1\!\!1_{\{X_T=b\}}\big]\,\mathrm{d}T = \int F(l)D_l^*\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\prod_{x\in\Lambda}\Big(\mathrm{d}l_x\,\frac{\mathrm{d}\theta_x}{2\pi}\Big). \tag{2.19}$$

Proof. Comparing (2.12) with this formula we see that it is enough to prove that the integration by parts formula

$$\int (D_l F)(l)\,\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\,\mathrm{d}^\Lambda l = \int F(l)\,\big(D_l^*\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\big)\,\mathrm{d}^\Lambda l \tag{2.20}$$

holds for any $\theta \in [0,2\pi]^\Lambda$. Since $D_l = \det_{ab}(-A-\partial_l)$ is a linear differential operator which is first order in each partial derivative, it suffices to consider one integral at a time and perform the integration by parts as follows: for any $x \in \Lambda$ and any fixed $(l_y)_{y\in\Lambda\setminus\{x\}}$,

$$\int_0^\infty \big(-\partial_{l_x}F(l)\big)\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\,\mathrm{d}l_x = \int_0^\infty F(l)\partial_{l_x}\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\,\mathrm{d}l_x, \qquad x \in \Lambda. \tag{2.21}$$

There are no boundary contributions because the map $l_x \mapsto F(l)$ has a compact support in $(0,\infty)$. This proves (2.19). $\square$

Step 4: For any $v \in \mathbb{C}^\Lambda$,

$$\int_0^\infty \mathbb{E}_a\big[\mathrm{e}^{\langle v,\ell_T\rangle}1\!\!1_{\{X_T=b\}}1\!\!1_{\{R_T=\Lambda\}}\big]\,\mathrm{d}T = \int \mathrm{e}^{\langle v,l\rangle}D_l^*\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\prod_{x\in\Lambda}\Big(\mathrm{d}l_x\,\frac{\mathrm{d}\theta_x}{2\pi}\Big). \tag{2.22}$$

Proof. Let $(f_n)_{n\in\mathbb{N}}$ be a uniformly bounded sequence of smooth functions with compact support in $(0,\infty)$ such that $f_n(t) \to 1\!\!1_{(0,\infty)}(t)$ for any t. Choose $F(l) = F_n(l) = \prod_{x\in\Lambda}(\mathrm{e}^{v_x l_x}f_n(l_x))$ in (2.19) and take the limit as $n \to \infty$, interchanging the limit with the integrals using the dominated convergence theorem. Observe that $\lim_{n\to\infty} F_n(\ell_T) = \mathrm{e}^{\langle v,\ell_T\rangle} \prod_{x\in\Lambda} \mathbb{1}_{(0,\infty)}(\ell_T(x)) = \mathrm{e}^{\langle v,\ell_T\rangle} \mathbb{1}_{\{R_T=\Lambda\}}$ almost surely. Furthermore, $\lim_{n\to\infty} F_n(l) = \mathrm{e}^{\langle v,l\rangle}$ almost everywhere with respect to the measure $\prod_{x\in\Lambda}(\mathrm{d}l_x \frac{\mathrm{d}\theta_x}{2\pi})$. Thus we obtain (2.22) in the limit of (2.19). □

Step 5: For all $v \in \mathbb{C}^\Lambda$,

$$\mathbb{E}_a\big[\mathrm{e}^{\langle v,\ell_T\rangle} \mathbb{1}_{\{X_T=b\}} \mathbb{1}_{\{R_T=\Lambda\}}\big] = \int_{\mathcal{M}_T^+(\Lambda)} \mathrm{e}^{\langle v,l\rangle} \rho_{ab}^{(\Lambda)}(l)\, \mathrm{d}^\Lambda l, \qquad T>0, a,b\in\Lambda, \tag{2.23}$$

where $\rho_{ab}^{(\Lambda)}(l)$ is given by (2.5).

Proof. Recall that $\sum_{x\in\Lambda} \ell_T(x) = T$ almost surely and that $\sum_{x\in\Lambda} l_x = T$ for $l \in \mathcal{M}_T^+(\Lambda)$. Hence, without loss of generality, we can assume that $\Re v \in (-\infty,0)^\Lambda$, since adding a constant $C\in\mathbb{R}$ to all the v_x results in adding a factor of e^{CT} on both sides. In (2.22) we replace v_x by $v_x-\lambda$ with $\lambda>0$. Then (2.22) becomes

$$\begin{aligned}\int_0^\infty \mathrm{e}^{-\lambda T}\mathbb{E}_a\big[\mathrm{e}^{\langle v,\ell_T\rangle} \mathbb{1}_{\{X_T=b\}} \mathbb{1}_{\{R_T=\Lambda\}}\big]\,\mathrm{d}T &= \int \mathrm{e}^{\langle v,l\rangle}\mathrm{e}^{-\lambda\sum_x l_x} D_l^* \mathrm{e}^{\langle\phi,A\overline{\phi}\rangle} \prod_{x\in\Lambda}\Big(\mathrm{d}l_x\,\frac{\mathrm{d}\theta_x}{2\pi}\Big)\\ &= \int_{(0,\infty)^\Lambda} \mathrm{e}^{\langle v,l\rangle}\mathrm{e}^{-\lambda\sum_x l_x} \rho_{ab}^{(\Lambda)}(l)\ \mathrm{d}^\Lambda l\\ &= \int_0^\infty \mathrm{e}^{-\lambda T}\Big[\int_{\mathcal{M}_T^+(\Lambda)} \mathrm{e}^{\langle v,l\rangle}\rho_{ab}^{(\Lambda)}(l)\,\sigma_T(\mathrm{d}l)\Big]\,\mathrm{d}T,\end{aligned} \tag{2.24}$$

where

$$\rho_{ab}^{(\Lambda)}(l) = \int_{[0,2\pi]^\Lambda} D_l^*\mathrm{e}^{\langle\phi,A\overline{\phi}\rangle}\prod_{x\in\Lambda}\frac{\mathrm{d}\theta_x}{2\pi}, \qquad l\in(0,\infty)^\Lambda. \tag{2.25}$$

In the second equation, we have interchanged the integrations over l and θ and have rewritten the θ integral using (2.5). In the third equation in (2.24), we have introduced the variable $T=\sum_x l_x$ and used (2.3).

Hence we have proved that the Laplace transforms with respect to T of the two sides of (2.23) coincide. As a consequence, (2.23) holds for almost every $T>0$. Furthermore, (2.23) even holds for all $T>0$, since both sides are continuous. Indeed, for small h we have $\langle v,\ell_{T+h}\rangle = \langle v,\ell_T\rangle$, $X_{T+h}=X_T$ and $R_{T+h}=R_T$ with high probability, which easily implies the continuity of the left hand side of (2.23). We see that the right hand side is continuous for $T>0$ by using the change of variable $t=T^{-1}l$ and (2.5) to rewrite the right hand side as an integral of a continuous function of T,t on the standard simplex $\mathcal{M}_{T=1}^+(\Lambda)$. □

Now we complete the proof of the theorem:

Step 6: The formula (2.4) holds for any finite or countably infinite state space Λ and any finite subset R of Λ.

Proof. It is enough to prove (2.4) for the case $F(l)=\mathrm{e}^{\langle v,l\rangle}$ with $\Re(v)\in(-\infty,0)^R$ because the distribution of $(\ell_T(x))_{x\in R}$ on the event $\{R_T=R\}$ is determined by its characteristic function.

Consider the Markov chain on R with conservative generator $A^{(R)}=(A^{(R)}_{x,y})_{x,y\in R}$ given by

$$A^{(R)}_{x,y} = \begin{cases} A_{x,y} & \text{if } x\neq y,\\ -\sum_{y\in R\setminus\{x\}} A_{x,y} & \text{if } x=y,\end{cases} \tag{2.26}$$

and let $V^{(R)}$ be the diagonal $R\times R$ matrix with $V^{(R)}_{x,x}=\sum_{y\in\Lambda\setminus R} A_{x,y}$. Then

$$A^{(R)}_{x,y} = A_{x,y} + V^{(R)}_{x,y} \qquad \forall x,y\in R. \tag{2.27}$$

When started in R, the Markov chain with generator $A^{(R)}$ coincides with the original one as long as no step to a site outside R is attempted. Step decisions outside R are suppressed. The distribution of this chain is absolutely continuous with respect to the original one. More precisely,

$$\mathbb{E}_a\Big[F(\ell_T)1\!\!1_{\{X_T=b\}}1\!\!1_{\{R_T=R\}}\Big] = \mathbb{E}_a^{(R)}\Big[F(\ell_T)\mathrm{e}^{-\sum_{x\in R}\ell_T(x)V_{x,x}^{(R)}}1\!\!1_{\{X_T=b\}}1\!\!1_{\{R_T=R\}}\Big], \qquad T>0, a,b\in R, \tag{2.28}$$

where $\mathbb{E}_a^{(R)}$ is the expectation with respect to the Markov chain on R with generator $A^{(R)}$. Applying (2.23) for this chain with $\mathrm{e}^{\langle v,l\rangle}$ replaced by

$$F_R(l) = F(l)\mathrm{e}^{-\sum_{x\in R} l_x V_{x,x}^{(R)}} \tag{2.29}$$

and with Λ replaced by R, we obtain, writing $\partial_l^{(R)}$ for the restriction of ∂_l to $R\times R$,

$$\begin{aligned}\mathbb{E}_a\Big[F(\ell_T)1\!\!1_{\{X_T=b\}}1\!\!1_{\{R_T=R\}}\Big] &= \mathbb{E}_a^{(R)}\Big[F_R(\ell_T)1\!\!1_{\{X_T=b\}}1\!\!1_{\{R_T=R\}}\Big]\\ &= \int_{\mathcal{M}_T^+(R)} F_R(l)\rho_{ab}^{(R)}(l)\,\sigma_T(\mathrm{d}l) = \int_{\mathcal{M}_T^+(R)} F(l)\widetilde{\rho}_{ab}^{(\Lambda,R)}(l)\,\sigma_T(\mathrm{d}l)\end{aligned} \tag{2.30}$$

where

$$\widetilde{\rho}_{ab}^{(\Lambda,R)}(l) = \mathrm{e}^{-\sum_{x\in R} l_x V_{x,x}^{(R)}} \det_{ab}(-A^{(R)}+\partial_l^{(R)}) \int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R}\phi_x A_{x,y}^{(R)}\overline{\phi}_y} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}. \tag{2.31}$$

By Lemma 2.4 followed by (2.27),

$$\begin{aligned}\widetilde{\rho}_{ab}^{(\Lambda,R)}(l) &= \det_{ab}(-A^{(R)}-V^{(R)}+\partial_l^{(R)})\Big[\mathrm{e}^{-\sum_{x\in R} l_x V_{x,x}^{(R)}}\int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R}\phi_x A_{x,y}^{(R)}\overline{\phi}_y} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}\Big]\\ &= \det_{ab}(-A^{(R)}-V^{(R)}+\partial_l^{(R)})\int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R}\phi_x A_{x,y}\overline{\phi}_y} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}\\ &= \det_{ab}^{(R)}(-A+\partial_l)\int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R}\phi_x A_{x,y}\overline{\phi}_y} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}.\end{aligned} \tag{2.32}$$

From the definition (2.5), and using (2.7), we recognise the last line as $\rho_{ab}^{(R)}(l)$. Therefore, by combining (2.32) and (2.30) we have proved (2.4) in the theorem. □

Now we collect some alternative expressions for the density $\rho_{ab}^{(R)}$:

Proposition 2.5. *Let the assumptions of Theorem 2.1 be satisfied. Let $B = ([1-\delta_{x,y}]A_{x,y})_{x,y\in\Lambda}$ be the off-diagonal part of A. Then, for any finite subset R of Λ and for any sites $a,b\in R$, and for any $l\in\mathcal{M}_T^+(R)$, the following holds:*

(i)

$$\rho_{ab}^{(R)}(l) = \mathrm{e}^{\sum_{x\in R} l_x A_{x,x}} \det_{ab}^{(R)}\big(-B+\partial_l\big)\int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R} B_{x,y}\sqrt{l_x}\sqrt{l_y}\mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}. \tag{2.33}$$

(ii) *For any $r\in(0,\infty)^R$,*

$$\rho_{ab}^{(R)}(l) = \mathrm{e}^{\sum_{x\in R} l_x A_{x,x}} \det_{ab}^{(R)}\big(-B+\partial_l\big)\int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R} r_x B_{x,y} r_y^{-1}\sqrt{l_x}\sqrt{l_y}\mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}. \tag{2.34}$$

(iii)

$$\rho_{ab}^{(R)}(l) = \int_{[0,2\pi]^R} \det_{ab}^{(R)}\big(-B+V_{\theta,l}\big)\mathrm{e}^{\sum_{x,y\in R} A_{x,y}\sqrt{l_x}\sqrt{l_y}\mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}, \tag{2.35}$$

where $V_{\theta,l} = (\delta_{x,y} v_{\theta,l}(x))_{x\in R}$ *is the diagonal matrix with entries*

$$v_{\theta,l}(x) = \sum_{z\in R} B_{x,z} \sqrt{\frac{l_z}{l_x}} \mathrm{e}^{\mathrm{i}(\theta_x - \theta_z)}, \qquad x \in R. \tag{2.36}$$

The formula in (2.34) will be helpful later when we derive upper bounds on $\rho_{ab}^{(R)}(l)$ in the case that A is not symmetric. The remainder of the paper does not rely on the formula in (2.35). However, we find (2.35) of independent interest, since the integral in (2.35) does not involve any derivative.

Proof. Formula (2.33) follows from (2.5) by using Lemma 2.4.

We now prove (2.34). Fix $r \in (0,\infty)^R$ and observe that, for any $l \in (0,\infty)^R$,

$$\int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R} B_{x,y} \sqrt{l_x}\sqrt{l_y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \prod_{x\in R} \frac{\mathrm{d}\theta_x}{2\pi} = \int_{[0,2\pi]^R} \mathrm{e}^{\sum_{x,y\in R} r_x B_{x,y} r_y^{-1} \sqrt{l_x}\sqrt{l_y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \prod_{x\in R} \frac{\mathrm{d}\theta_x}{2\pi}. \tag{2.37}$$

Indeed, substituting $\mathrm{e}^{\mathrm{i}\theta_x} = z_x$ for $x \in R$, we can rewrite the integrals as integrals over circles in the complex plane. The integrand is analytic in $z_x \in \mathbb{C} \setminus \{0\}$. Hence, the integral is independent of the curve (as long as it is closed and winds around zero precisely once), and it is equal to the integral along the centred circle with radius r_x instead of radius one. Re-substituting $r_x \mathrm{e}^{\mathrm{i}\theta_x} = z_x$, we arrive at (2.37). Comparing to (2.33), we see that we have derived (2.34).

Finally, we prove (2.35). We use (2.34) with $r = \sqrt{l}$ and interchange $\det_{ab}^{(R)}\big(-B+\partial_l\big)$ with $\int_{[0,2\pi]^R}$ (this is justified by the analyticity of the integrand in all the l_x with $x \in R$). This gives that

$$\rho_{ab}^{(R)}(l) = \mathrm{e}^{\sum_{x\in R} l_x A_{x,x}} \int_{[0,2\pi]^R} \det{}_{ab}^{(R)}\big(-B+\partial_l\big) \mathrm{e}^{\sum_{x,y\in R} l_x B_{x,y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \prod_{x\in R} \frac{\mathrm{d}\theta_x}{2\pi}.$$

Use Lemma 2.4 with $g = 1$ to see that

$$\det{}_{ab}^{(R)}\big(-B+\partial_l\big) \mathrm{e}^{\sum_{x,y\in R} l_x B_{x,y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} = \mathrm{e}^{\sum_{x,y\in R} l_x B_{x,y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}} \det{}_{ab}^{(R)}\big(-B+\widetilde{V}_\theta\big),$$

where $\widetilde{V}_\theta = (\delta_{x,y}\widetilde{v}_\theta(x))_{x\in R}$ is the diagonal matrix with entries $\widetilde{v}_\theta(x) = \sum_{z\in R} B_{x,z} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_z)}$.

Now we use the same transformation as in (2.37): We interpret the integrals over θ_x as integrals over circles of radius $\sqrt{l_x}$ and replace them by integrals over circles with radius one. By this transformation, $\widetilde{V}_\theta$ is transformed into $V_{\theta,l}$, and the term $\mathrm{e}^{\sum_{x,y\in R} l_x B_{x,y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}}$ is transformed into $\mathrm{e}^{\sum_{x,y\in R} \sqrt{l_x} B_{x,y} \sqrt{l_y} \mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}}$. Recalling that B is the off-diagonal part of A, (2.35) follows. $\square$

3. Large deviation upper bounds for the local times

In this section we use Theorem 2.1 to derive sharp upper bounds for the probability in (1.2) and for the expectation in (1.5) for fixed T and fixed finite ranges of the local times. The main term in this estimate is given in terms of the rate function I_A. The main value of our formula, however, comes from the facts that (1) the error term is controlled on a subexponential scale, (2) the set Γ in (1.2) is just assumed measurable, and (3) the functional F in (1.5) is just assumed measurable. Let us stress that this formula is extremely useful, since the functional F is not upper semicontinuous nor bounded in many important applications.

In Section 3.1 we give a pointwise upper bound for the density, in Section 3.2 we apply it to derive upper bounds for the probability in (1.2) and for the expectation in (1.5), and in Section 3.3 we consider the same problem for state spaces $\Lambda = \Lambda_T \subseteq \mathbb{Z}^d$ depending on T and increasing to $\mathbb{Z}^d$.

3.1. **Pointwise upper bound for the density.** Here is a pointwise upper bound for the density. Recall the rate function I_A introduced in (1.4).

Proposition 3.1 (Upper bound for $\rho_{ab}^{(R)}$). *Under the assumptions of Theorem 2.1, for any finite subset R of Λ, and for any $a,b\in R$, any $T>0$ and any $l\in\mathcal{M}_T^+(R)$,*

$$\rho_{ab}^{(R)}(l)\leq \mathrm{e}^{-TI_A(\frac{1}{T}l)}\Big(\prod_{x\in R\setminus\{a,b\}}\sqrt{\frac{T}{l_x}}\Big)\eta_R^{|R|-1}\mathrm{e}^{[\eta_R^{-1}+(4\eta_R^2T)^{-1}]\sum_{x,y\in R}\sqrt{l_x}g_yB_{x,y}/(\sqrt{l_y}g_x)},\tag{3.1}$$

where $g\in(0,\infty)^R$ is the minimizer in (1.4) *for $\mu=\frac{l}{T}$ and*

$$\eta_R=\max\Big\{\max_{x\in R}\sum_{y\in R\setminus\{x\}}|B_{x,y}|,\ \max_{y\in R}\sum_{x\in R\setminus\{y\}}|B_{x,y}|,\ 1\Big\},\tag{3.2}$$

where $B=([1-\delta_{x,y}]A_{x,y})_{x,y\in\Lambda}$ is the off-diagonal part of A.

Remark 3.2. If A (and hence B) is symmetric, then $g=\sqrt{l}$ is the minimizer in (1.4), and we have $I_A(\mu)=\|(-A)^{\frac12}\sqrt{\mu}\|_2^2$. In this case the upper bound simplifies to

$$\rho_{ab}^{(R)}(l)\leq \mathrm{e}^{-TI_A(\frac{1}{T}l)}\Big(\prod_{x\in R\setminus\{a,b\}}\sqrt{\frac{T}{l_x}}\Big)\eta_R^{|R|-1}\mathrm{e}^{|R|[1+(4\eta_RT)^{-1}]}.\tag{3.3}$$

$\Diamond$

The proof of Proposition 3.1 makes use of three lemmas that we will state and prove first.

Lemma 3.3. *Let $\widetilde{B}\in[0,\infty)^{R\times R}$ be any matrix with nonnegative elements, and let $Q\subseteq R$. Then*

$$0\leq\partial_l^Q\int_{[0,2\pi]^R}\mathrm{e}^{\sum_{x,y\in R}\widetilde{B}_{x,y}\sqrt{l_x}\sqrt{l_y}\mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}}\prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}\leq\partial_l^Q\mathrm{e}^{\sum_{x,y\in R}\widetilde{B}_{x,y}\sqrt{l_x}\sqrt{l_y}},\qquad l\in(0,\infty)^R,\tag{3.4}$$

where $\partial_l^Q=\prod_{x\in Q}\partial_{l_x}$.

Proof. Write $\mathrm{e}^{\sum_{x,y\in R}\cdots}=\prod_{x,y\in R}\mathrm{e}^{\cdots}$ and expand the exponentials as power series. For $n=(n_{x,y})_{x,y\in R}\in\mathbb{N}_0^{R\times R}$, we write $n!=\prod_{x,y\in R}n_{x,y}!$. Then we obtain

$$\begin{aligned}&\partial_l^Q\int_{[0,2\pi]^R}\mathrm{e}^{\sum_{x,y\in R}\widetilde{B}_{x,y}\sqrt{l_x}\sqrt{l_y}\mathrm{e}^{\mathrm{i}(\theta_x-\theta_y)}}\prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}\\&\quad=\sum_{n\in\mathbb{N}_0^{R\times R}}\frac{1}{n!}\partial_l^Q\Big[\prod_{x,y\in R}\big(\widetilde{B}_{x,y}\sqrt{l_x}\sqrt{l_y}\big)^{n_{x,y}}\int_{[0,2\pi]^R}\mathrm{e}^{\mathrm{i}\sum_{x,y\in R}n_{x,y}(\theta_x-\theta_y)}\prod_{x\in R}\frac{\mathrm{d}\theta_x}{2\pi}\Big].\end{aligned}\tag{3.5}$$

After rewriting the exponent in the integral on the right hand side using $\sum_{x,y}n_{x,y}(\theta_x-\theta_y)=\sum_x n_x\theta_x$, where $n_x=\sum_y(n_{x,y}-n_{y,x})$, it is clear that the integral equals one or zero. Hence, the lower bound in (3.4) is clear, and the upper bound comes from replacing the integral by one and a resummation over n. $\Box$

Lemma 3.4. *Fix any matrix $B\in\mathbb{R}^{R\times R}$, let $a,b\in R$, and let $f\colon(0,\infty)^R\to\mathbb{R}$ be any function with nonnegative derivatives, i.e., $\partial_l^Qf(l)\geq0$ for all $Q\subseteq R$. Then*

$$\big|\det\nolimits_{a,b}^{(R)}(-B+\partial_l)f\big|\leq\eta_R\prod_{x\in R\setminus\{a,b\}}\big(\eta_R+\partial_{l_x}\big)f,\tag{3.6}$$

where η_R is defined in (3.2).

Proof. Recalling that the determinant is the (signed) volume subtended by the rows, we can bound a determinant by the product of the lengths of the rows. This is called the Hadamard bound and it applies to any real square matrix. Therefore, for $X\subseteq R$ and $a,b\in X$,

$$\big|\det\nolimits_{ab}^{(X)}(-B)\big|\leq\prod_{x\in X\setminus\{b\}}\|B_x\|\leq\prod_{x\in X\setminus\{b\}}\eta_R=\eta_R^{|X|-1},$$

where B_x is the row x of B after eliminating the a-th column, and $\|\cdot\|$ is the Euclidean length, which is bounded by η_R because $\sum|a_i|^2 \le (\sum|a_i|)^2$. Also,

$$\det\nolimits_{ab}^{(R)}(-B+\partial_l)f(l) = \sum_{\sigma\colon R\setminus\{b\}\to R\setminus\{a\}} \operatorname{sign}(\hat\sigma) \prod_{x\in R\setminus\{a\}} \big(-B_{x,\sigma_x}+\delta_{x,\sigma_x}\partial_{l_x}\big)f(l), \tag{3.7}$$

where the sum over σ is over all bijections $R\setminus\{b\}\to R\setminus\{a\}$, and where $\operatorname{sign}(\hat\sigma)$ is the sign of the permutation $\hat\sigma\colon R\mapsto R$ obtained by letting $\hat\sigma_x=\sigma_x$ for $x\neq b$ and $\hat\sigma_b=a$. Expanding the product, we obtain

$$\begin{aligned}\det\nolimits_{ab}^{(R)}(-B+\partial_l)f(l) &= \sum_{Q\subseteq R\setminus\{a,b\}} \sum_{\sigma\colon Q^{\mathrm c}\setminus\{b\}\to Q^{\mathrm c}\setminus\{a\}} \operatorname{sign}(\hat\sigma)\Big(\prod_{x\in Q^{\mathrm c}\setminus\{b\}}\big(-B_{x,\sigma_x}\big)\Big)\Big(\prod_{x\in Q}\partial_{l_x}\Big)f(l)\\ &= \sum_{Q\subseteq R\setminus\{a,b\}} \det\nolimits_{ab}^{(Q^{\mathrm c})}(-B)\Big(\prod_{x\in Q}\partial_{l_x}\Big)f(l),\end{aligned} \tag{3.8}$$

where we write $Q^{\mathrm c}=R\setminus Q$. Take absolute values and bound the cofactor using the Hadamard bound,

$$\begin{aligned}\big|\det\nolimits_{ab}^{(R)}(-B+\partial_l)f(l)\big| &\le \sum_{Q\subseteq R\setminus\{a,b\}} \eta_R^{|Q^{\mathrm c}\setminus\{b\}|}\Big(\prod_{x\in Q}\partial_{l_x}\Big)f(l)\\ &= \eta_R \sum_{Q\subseteq R\setminus\{a,b\}}\Big(\prod_{x\in(R\setminus\{a,b\})\setminus Q}\eta_R\Big)\Big(\prod_{x\in Q}\partial_{l_x}\Big)f(l)\\ &= \eta_R\prod_{x\in R\setminus\{a,b\}}\big(\eta_R+\partial_{l_x}\big)f(l).\end{aligned} \tag{3.9}$$

□

Lemma 3.5. *Fix any finite subset R of Λ, let $\widetilde B\in[0,\infty)^{R\times R}$ be any matrix with nonnegative elements, and fix $a,b\in R$. Then, for any $T>0$ and any $l\in\mathcal M_T^+$,*

$$\begin{aligned}&\det\nolimits_{ab}^{(R)}\big(-B+\partial_l\big)\int_{[0,2\pi]^R} \mathrm e^{\sum_{x,y\in R}\widetilde B_{x,y}\sqrt{l_x}\sqrt{l_y}\mathrm e^{\mathrm i(\theta_x-\theta_y)}}\prod_{x\in R}\frac{\mathrm d\theta_x}{2\pi}\\ &\le \mathrm e^{\sum_{x,y\in R}\widetilde B_{x,x}\sqrt{l_x}\sqrt{l_y}}\Big(\prod_{x\in R\setminus\{a,b\}}\sqrt{\frac{T}{l_x}}\Big)\eta_R^{|R|-1}\mathrm e^{[\eta_R{}^{-1}+(4\eta_R^2T)^{-1}]\sum_{x,y\in R}\widetilde B_{x,y}},\end{aligned} \tag{3.10}$$

where η_R is defined in (3.2).

Proof. By Lemma 3.4 followed by Lemma 3.3, we obtain

$$\text{l.h.s. of (3.10)} \le \eta_R\prod_{x\in R\setminus\{a,b\}}\big(\eta_R+\partial_{l_x}\big)\mathrm e^{\sum_{x,y\in R}\widetilde B_{x,y}\sqrt{l_x}\sqrt{l_y}}. \tag{3.11}$$

Substitute $t_x=\frac{\sqrt{l_x}}{\sqrt T}\in[0,1]$ and abbreviate $f(t)=\mathrm e^{T\sum_{x,y\in R}\widetilde B_{x,y}t_xt_y}$. By the chain rule, $\partial_{l_x}=\frac{1}{2T}\frac{1}{t_x}\partial_{t_x}$. Then

$$\begin{aligned}\text{l.h.s. of (3.10)} &\le \eta_R^{|R|-1}\prod_{x\in R\setminus\{a,b\}}\Big(1+\frac{1}{2\eta_RT}\frac{1}{t_x}\partial_{t_x}\Big)f(t)\\ &\le \eta_R^{|R|-1}\Big(\prod_{x\in R\setminus\{a,b\}}\frac{1}{t_x}\Big)\prod_{x\in R\setminus\{a,b\}}\Big(1+\frac{1}{2\eta_RT}\partial_{t_x}\Big)f(t),\end{aligned} \tag{3.12}$$

where we have used that $t_x \leq 1$. Since *all* t derivatives (not just the first order derivatives) of f are nonnegative since $\widetilde{B}_{x,y} \geq 0$, we can add in some extra derivatives and continue the bound with

$$\begin{aligned} \text{l.h.s. of (3.10)} &\leq \eta_R^{|R|-1}\Big(\prod_{x\in R\setminus\{a,b\}}\frac{1}{t_x}\Big)\prod_{x\in R\setminus\{a,b\}}\Big(\sum_{n=0}^\infty\frac{1}{n!}\frac{\partial_{t_x}^n}{(2\eta_R T)^n}\Big)f(t) \\ &= \eta_R^{|R|-1}\Big(\prod_{x\in R\setminus\{a,b\}}\frac{1}{t_x}\Big)\, f\Big(t+(2\eta_R T)^{-1}\mathbb{1}_R\Big), \end{aligned} \tag{3.13}$$

where the last equation follows from Taylor's theorem, and $\mathbb{1}_R\colon R\to\{1\}$ is the constant function.

Recalling that $t_x \leq 1$, we may estimate

$$\begin{aligned} \frac{1}{T}\log f\Big(t+(2\eta_R T)^{-1})\mathbb{1}_R\Big) &= \sum_{x,y\in R}\widetilde{B}_{x,y}t_xt_y + \frac{1}{2\eta_R T}\sum_{x,y\in R}\widetilde{B}_{x,y}(t_x+t_y) + \frac{1}{(2\eta_R T)^2}\sum_{x,y\in R}\widetilde{B}_{x,y} \\ &\leq \frac{1}{T}\log f(t) + \frac{1}{T}\Big[\frac{1}{\eta_R}+\frac{1}{4\eta_R^2 T}\Big]\sum_{x,y\in R}\widetilde{B}_{x,y}. \end{aligned}$$

We conclude that

$$\text{l.h.s. of (3.10)} \leq \eta_R^{|R|-1}\Big(\prod_{x\in R\setminus\{a,b\}}\frac{1}{t_x}\Big)\, f(t)\,\mathrm{e}^{[\eta_R^{-1}+(4\eta_R^2 T)^{-1}]\sum_{x,y\in R}\widetilde{B}_{x,y}}. \tag{3.14}$$

Re-substituting $t_x = \sqrt{l_x/T}$ and $f(t) = \mathrm{e}^{T\sum_{x,y\in R}\widetilde{B}_{x,y}t_xt_y}$, the lemma is proved. □

Proof of Proposition 3.1. Fix any $r\in(0,\infty)$ and recall the representation of the density $\rho_{ab}^{(R)}$ in (2.34). Now apply Lemma 3.5 for $\widetilde{B} = (r_xB_{x,y}r_y^{-1})_{x,y\in R}$, to obtain

$$\rho_{ab}^{(R)}(l) \leq \mathrm{e}^{\sum_{x,y\in R}r_x\sqrt{l_x}A_{x,y}\sqrt{l_y}r_y^{-1}}\Big(\prod_{x\in R\setminus\{a,b\}}\sqrt{\frac{T}{l_x}}\Big)\eta_R^{|R|-1}\mathrm{e}^{[\eta_R^{-1}+(4\eta_R^2T)^{-1}]\sum_{x,y\in R}\widetilde{B}_{x,y}}.$$

Now we choose $r = \sqrt{l}/g$, where $g\in(0,\infty)^R$ is the minimiser in (1.4) for $\mu = \frac{1}{T}l$. This implies the bound in (3.1). □

3.2. **Upper bounds in the LDP and in Varadhan's lemma.** In this section we specialize to Markov chains having a symmetric generator A and give a simple upper bound for the left hand side of (1.2) and for the expectation in (1.5). Recall from the text below (1.4) that, in the present case of a symmetric generator, $I_A(\mu) = \|(-A)^{\frac{1}{2}}\sqrt{\mu}\|_2^2$ for any probability measure μ on Λ.

Theorem 3.6 (Large deviation upper bounds for the local times)**.** *Let the assumptions of Theorem 2.1 be satisfied. Assume that A is symmetric. Fix a finite subset S of Λ. Then, for any $T\geq 1$ and any $a\in S$, with η_S as in (3.2), the following bounds hold:*

(i) *For every measurable $\Gamma\subseteq\mathcal{M}_1(S)$,*

$$\log\mathbb{P}_a\big(\tfrac{1}{T}\ell_T\in\Gamma, R_T\subseteq S\big) \leq -T\inf_{\mu\in\Gamma}\big\|(-A)^{\frac{1}{2}}\sqrt{\mu}\big\|_2^2 + |S|\log\big(\eta_S\sqrt{8\mathrm{e}T}\big) + \log|S| + \frac{|S|}{4T}. \tag{3.15}$$

(ii) *For every measurable functional $F\colon\mathcal{M}_1(S)\to\mathbb{R}$,*

$$\log\mathbb{E}_a\big[\mathrm{e}^{TF(\frac{1}{T}\ell_T)}\mathbb{1}_{\{R_T\subseteq S\}}\big] \leq T\sup_{\mu\in\mathcal{M}_1(S)}\Big[F(\mu)-\big\|(-A)^{\frac{1}{2}}\sqrt{\mu}\big\|_2^2\Big]+|S|\log\big(\eta_S\sqrt{8\mathrm{e}T}\big)+\log|S|+\frac{|S|}{4T}. \tag{3.16}$$

Theorem 3.6 is a significant improvement over the standard estimates known in large deviation theory. In fact, one standard technique to derive upper bounds for the left hand side of (3.15) is the use of the exponential Chebyshev inequality and a compactness argument if Γ is assumed closed. One important ingredient there is a good control on the logarithmic asymptotics of the expectation in (3.16) for linear functions F. This technique produces an error of order $\mathrm{e}^{o(T)}$, which can in general not be controlled on a smaller scale.

The standard technique to derive improved bounds on the expectation in (3.16) for fixed T is restricted to *linear* functions F, say $F(\cdot) = \langle V, \cdot\rangle$. This technique goes via an eigenvalue expansion for the operator $A + V$ in the set S with zero boundary condition. The main steps are the use of the Rayleigh-Ritz principle for the identification of the principal eigenvalue, and Parseval's identity. This gives basically the same result as in (3.16), but is strictly limited to linear functions F.

Proof of Theorem 3.6. It is clear that (ii) follows from (i), hence we only prove (i).

According to Theorem 2.1, we may express the probability on the left hand side of (3.15) as

$$\mathbb{P}_a\big(\tfrac{1}{T}\ell_T \in \Gamma, R_T \subseteq S\big) = \sum_{b\in S}\sum_{R\subseteq S\colon a,b\in R}\int_{\mathcal{M}_T^+(R)\cap\Gamma_{T,R}} \rho_{ab}^{(R)}(l)\,\sigma_T(\mathrm{d}l), \tag{3.17}$$

where $\Gamma_{T,R} = T\Gamma_R$, and Γ_R is the set of the restrictions of all the elements of Γ to R.

We fix $a, b \in S$ and $R \subseteq S$ with $a, b \in R$ and use the bound in Proposition 3.1, more precisely, the one in (3.3). Hence, for $l \in \mathcal{M}_T^+(R)\cap\Gamma_{T,R}$, we obtain, after a substitution $l = T\mu$ in the exponent, that

$$\rho_{ab}^{(R)}(l) \le \mathrm{e}^{-T\inf_{\mu\in\Gamma\colon \mathrm{supp}(\mu)\subseteq R}\|(-A)^{\frac12}\sqrt{\mu}\|_2^2}\Big(\prod_{x\in R\setminus\{a,b\}}\sqrt{\frac{T}{l_x}}\Big)\eta_R^{|R|-1}\mathrm{e}^{|R|[1+(4\eta_R T)^{-1}]}. \tag{3.18}$$

Substituting this in (3.17) and integrating over $l \in \mathcal{M}_T^+(R)$, we obtain

$$\begin{aligned}\mathbb{P}_a\big(\tfrac{1}{T}\ell_T \in \Gamma, R_T \subseteq S\big) &\le \mathrm{e}^{-T\inf_{\mu\in\Gamma}\|(-A)^{\frac12}\sqrt{\mu}\|_2^2}\eta_R^{|R|-1}\mathrm{e}^{|R|[1+(4\eta_R T)^{-1}]}\\ &\qquad\times\sum_{b\in S}\sum_{R\subseteq S\colon a,b\in R}\int_{\mathcal{M}_T^+(R)}\prod_{x\in R\setminus\{a\}}\sqrt{\frac{T}{l_x}}\,\sigma_T(\mathrm{d}l) \qquad (3.19)\\ &\le \mathrm{e}^{-T\inf_{\mu\in\Gamma}\|(-A)^{\frac12}\sqrt{\mu}\|_2^2}\eta_R^{|R|-1}\mathrm{e}^{|S|[1+(4\eta_R T)^{-1}]}|S|\sqrt{8}^{|S|}T^{|S|-1}.\end{aligned}$$

In the last integral, we have eliminated $l_a = T - \sum_{y\in R\setminus\{a\}} l_y$, have extended the $(|R|-1)$ single integration areas to $(0, T)$ and used that $\int_0^T l_x^{-\frac12}\,\mathrm{d}l_x = \sqrt{2T}$. Now we use that η_R is increasing in R and greater than or equal to one to arrive at (3.15). This completes the proof of (i). □

3.3. **Rescaled local times.** As an application of Theorem 3.6, we now consider continuous-time simple random walk restricted to a large T-dependent subset $\Lambda = \Lambda_T$ of $\mathbb{Z}^d$ increasing to $\mathbb{Z}^d$. We derive the sharp upper bound in the large deviation principle for its *rescaled* local times. Assume, for some scale function $T \mapsto \alpha_T \in (0,\infty)$, that Λ_T is equal to the box $[-R\alpha_T, R\alpha_T]^d \cap \mathbb{Z}^d$, where the scale function α_T satisfies

$$1 \ll \alpha_T \ll \Big(\frac{T}{\log T}\Big)^{\frac{1}{d+2}} \qquad \text{as } T\to\infty. \tag{3.20}$$

We introduce the rescaled version of the local times,

$$L_T(x) = \frac{\alpha_T^d}{T}\ell_T\big(\lfloor \alpha_T x\rfloor\big), \qquad x\in\mathbb{R}^d.$$

Note that L_T is a random step function on $\mathbb{R}^d$. In fact, it is a random probability density on $\mathbb{R}^d$. Its support is contained in the cube $[-R, R]^d$ if and only if the support of ℓ_T is contained in the box $[-R\alpha_T, R\alpha_T]^d \cap \mathbb{Z}^d$.

It is known that, as $T \to \infty$, the family $(L_T)_{T>0}$ satisfies a large deviation principle under the sub-probability measures $\mathbb{P}(\cdot \cap \{\mathrm{supp}(L_T) \subseteq [-R, R]^d\})$ for any $R > 0$. The speed is $T\alpha_T^{-2}$, and the rate function is the energy functional, i.e., the map $g^2 \mapsto \frac12\|\nabla g\|_2^2$, restricted to the set of squares g^2 of L^2-normalized functions g such that g lies in $H^1(\mathbb{R}^d)$ and has its support in $[-R, R]^d$. The topology is the one which is induced by all the test integrals of g^2 against continuous and bounded functions. This large-deviation principle is proved in [GKS05] for the discrete-time random walk, and the proof for continuous-time walks is rather similar (see also [HKM05], where the proof of this fact is sketched). Hence, Varadhan's lemma yields precise logarithmic asymptotics for all exponential functionals of L_T that are bounded and continuous in the above mentioned topology.

Note that this large deviations principle for L_T is almost the same as the one which is satisfied by the normalized Brownian occupation times measures (see [Gär77, DV75-83]), the main difference being the speed (which is T in [Gär77, DV75-83] instead of $T\alpha_T^{-2}$ here) and the fact that L_T does not take values in the set of continuous functions $\mathbb{R}^d \to [0,\infty)$.

Here we want to point out that Theorem 3.6 yields a new method to derive upper bounds for many exponential functionals of L_T. For a cube $Q\subset \mathbb{R}^d$, we denote by $M_1(Q)$ the set of all probability densities $Q \to [0,\infty)$.

Theorem 3.7. *Fix $R > 0$, denote $Q_R = [-R,R]^d$ and fix a measurable function $F\colon M_1(Q_R) \to \mathbb{R}$. Introduce*

$$\chi = \inf\Big\{\frac{1}{2}\|\nabla g\|_2^2 - F(g^2)\colon g \in H^1(\mathbb{R}^d), \|g\|_2 = 1, \mathrm{supp}(g) \subseteq Q_R\Big\}. \tag{3.21}$$

Then

$$\limsup_{T\to\infty} \frac{\alpha_T^2}{T} \log \mathbb{E}\Big[\exp\Big\{\frac{T}{\alpha_T^2} F(L_T)\Big\} 1\!\!1_{\{\mathrm{supp}(L_T)\subseteq Q_R\}}\Big] \le -\chi, \tag{3.22}$$

provided that

$$\liminf_{T\uparrow\infty} \inf_{\mu\in\mathcal{M}_1(B_{R\alpha_T})} \Big(\alpha_T^2 \frac{1}{2}\sum_{x\sim y}\Big(\sqrt{\mu(x)}-\sqrt{\mu(y)}\Big)^2 - F\Big(\alpha_T^d \mu\big(\lfloor\cdot\,\alpha_T\rfloor\big)\Big)\Big) \ge \chi. \tag{3.23}$$

Proof. Introduce

$$F_T(\mu) = \frac{1}{\alpha_T^2} F\Big(\alpha_T^d \mu\big(\lfloor\cdot\,\alpha_T\rfloor\big)\Big), \qquad \mu \in \mathcal{M}_1(\mathbb{Z}^d),$$

then we have $\frac{1}{\alpha_T^2}F(L_T) = F_T(\frac{1}{T}\ell_T)$. Hence, Theorem 3.6(ii) yields that

$$\begin{aligned}\mathbb{E}\Big[\exp\Big\{\frac{T}{\alpha_T^2}F(L_T)\Big\}1\!\!1_{\{\mathrm{supp}(L_T)\subseteq Q_R\}}\Big] &= \mathbb{E}\Big[\exp\Big\{T\,F_T(\tfrac{1}{T}\ell_T)\Big\}1\!\!1_{\{\mathrm{supp}(\ell_T)\subseteq Q_{R\alpha_T}\}}\Big]\\ &\le \mathrm{e}^{o(T\alpha_T^{-2})}\mathrm{e}^{-T\chi_T},\end{aligned}$$

where

$$\chi_T = \inf_{\mu\in\mathcal{M}_1(Q_{R\alpha_T}\cap\mathbb{Z}^d)} \Big(\frac{1}{2}\sum_{x\sim y}\Big(\sqrt{\mu(x)}-\sqrt{\mu(y)}\Big)^2 - F_T(\mu)\Big).$$

Here we used that the two error terms on the right hand side of (3.16) are $\mathrm{e}^{o(T\alpha_T^{-2})}$ since $\eta_S \le 2d$ for any $S \subseteq \mathbb{Z}^d$ and because of our growth assumption in (3.20). Now (3.22) follows from (3.23). □

Theorem 3.7 proved extremely useful in the study of the parabolic Anderson model in [HKM05]. Indeed, it was crucial in that paper to find the precise upper bound of the left hand side of (3.22) for the functional

$$F(g^2) = \int_{Q_R} g^2(x)\log g^2(x)\,\mathrm{d}x,$$

which has bad continuity properties in the topology in which the above mentioned large deviations principle holds. However, Theorem 3.7 turned out to be applicable since the crucial prerequisite in (3.23) had been earlier provided in [GH99]. The main methods there were equicontinuity, uniform integrability and Arzela-Ascoli's theorem.

In the same paper [HKM05], also the functional

$$F(g^2) = -\int_{Q_R} |g(x)|^{2\gamma}\,\mathrm{d}x, \qquad \text{with some } \gamma \in (0,1),$$

was considered. This problem arose in the study of the parabolic Anderson model for another type of potential distribution which was earlier studied in [BK01]. The prerequisite in (3.23) was provided in [HKM05] using techniques from Gamma-convergence; see [AC04] for these techniques.

4. Discussion

In this section, we give some comments on the history of the problem addressed in the present paper.

4.1. **Historical background.** The formulas in this paper have been motivated by the work of the theoretical physicist J.M. Luttinger [Lut83] who gave a (non-rigorous) asymptotic evaluation of certain path integrals. Luttinger claimed that there is an asymptotic series

$$\mathbb{E}\big[\mathrm{e}^{-TF(\ell_T/T)}\big] \sim \sqrt{T}\mathrm{e}^{-c_0 T}\Big(c_1 + \frac{c_2}{T} + \frac{c_3}{T^2} + \cdots\Big)$$

for Brownian local times. He provided an algorithm to compute all the coefficients. He showed that his algorithm gives the Donsker-Varadhan large deviations formula for c_0 and he explicitly computed the central limit correction c_1.

In [BM91] Brydges and Muñoz-Maya used Luttinger's methods to verify that his asymptotic expansion is valid to all orders for a Markov process with symmetric generator and finite state space. The hypotheses are that F is smooth and the variational principle that gives the large deviations coefficient c_0 is non-degenerate. Luttinger implicitly relies on similar assumptions when he uses the Feynman expansion for his functional integral.

Thus there remains the open problem to prove that Luttinger's series is asymptotic for more general state spaces, in particular, for Brownian motion. As far as we know, the best progress to date is in [BDT95] where compact state spaces were considered and the asymptotics including the c_1 correction was verified.

Luttinger's paper used a calculus called *Grassman integration*. The background to this is that the Feynman-Kac formula provides a probabilistic representation for the propagation of elementary particles that satisfy "Bose statistics". To obtain a similar representation for elementary particles that satisfy "Fermi statistics" one is led in [Ber87] to an analogue of integration defined as a linear functional on a non-Abelian Grassman algebra in place of the Abelian algebra of measurable functions: this is Grassman integration. An important part of this line of thought concerns a case where there is a relation called *supersymmetry*. This background gives no hint that Grassman integrals are relevant for ordinary Markov processes, but, nevertheless, Parisi and Sourlas [PS80] and McKane [McK80] noted that random walk expectations can be expressed in terms of the Grassman extension of Gaussian integration. Luttinger followed up on these papers by being much more explicit and precise about the supersymmetric representation in terms of Grassman integration and by deriving his series.

In [LeJ87] Le Jan pointed out that Grassman integration in this context is actually just ordinary integration in the context of differential forms. The differential forms are the non-Abelian algebra and the standard definition of integration of differential forms provides the linear functional. Since integration over differential forms is defined in terms of ordinary integration one can remove the differential forms, as we have done in this paper, but this obscures the underlying mechanism of supersymmetry. The formalism with differential forms is explained in [BI03b, page 551] where it is used to study the Green's function of a self-repelling walk on a hierarchical lattice. Two other applications of the same formalism are the proof of the Matrix-Tree theorem in [Abd03] and a result on self-avoiding trees given in [BI03a].

Luttinger found an instance of a relation between the local time of a Markov process on a state space E and the square of a *Gaussian field* indexed by E. The first appearance of such a relation was given by Symanzik in [Sym69]. His statement is that the sum of the local times of an ensemble of Brownian loops is the square of a Gaussian field. The references given above to Parisi-Sourlas, McKane and

Luttinger removed the need for an ensemble by bringing, in its place, Grassman integration. The paper of Symanzik was not immediately rigorous because he claimed his result for Brownian motion but it makes almost immediate sense for Markov processes on finite state spaces only. Based on this work a rigorous relation between the square of a Gaussian field and local time of a random walk on a lattice was given by Brydges, Fröhlich and Spencer in [BFS82]. Dynkin [Dyn83, Dyn84b, Dyn84a] showed that the identities of that paper can be extended to Brownian motion in one and two dimensions. In this form, the *Dynkin Isomorphism*, it became a useful tool for studying local time of diffusions and much work has been done by Rosen and Marcus in exploiting and extending these ideas, e.g., see [MR96, EKMRS00]. The relation between the local time and the square of a Gaussian field is concealed in this paper in (2.12) which relates the local time ℓ to $l = |\phi|^2$ where ϕ is Gaussian. This is more obvious when ϕ is expressed as $\phi = u + \mathrm{i}v$ instead of in terms of polar coordinates $\phi = \sqrt{l}\mathrm{e}^{\mathrm{i}\theta}$.

4.2. **Relation to the Ray-Knight theorem.** Our density formula in Theorem 2.1 can also be used to prove a version of the *Ray-Knight theorem* for continous-time simple random walk on $\mathbb{Z}$. The well-known Ray-Knight theorem for one-dimensional Brownian motion, see [RY91, Sections XI.1-2], [KS91, Sections 6.3-4], was originally proved in [Kni63, Ray63]. It describes the Brownian local times, observed at certain stopping times, as a homogeneous Markov chain in the spatial parameter. Numerous deeper investigations of this idea have been made, e.g., for general symmetric Markov processes [EKMRS00], for diffusions with fixed birth and death points on planar cycle-free graphs [EK93, EK96], and on the relations to Dynkin's isomorphism [She85], [Eise94].

The (time and space) discrete version of the Ray-Knight theorem, i.e., for simple random walk on $\mathbb{Z}$, was also introduced in [Kni63], however it turned out there that it is not the local times on the sites, but on the edges that enjoys a Markov property. This idea has been used or re-invented a couple of times, e.g., for applications to random walk in random environment [KKS75], to reinforced random walk [Toth96], and to random polymer measures [GH93].

In the present situation of continuous time and discrete space, it turns out that the local times themselves form a nice Markov chain. However, a proof appears to be missing. In fact, up to our best knowledge, [MS87] is the only paper that provides (the outline of) a proof, but only for the special case where the walk starts and ends in the same point.

For fixed $b \in \mathbb{Z}$, we denote

$$T_b^h = \inf\{t > 0 \colon \ell_t(b) > h\}, \qquad h > 0, \tag{4.1}$$

the right-continuous inverse of the map $t \mapsto \ell_t(b)$. We denote by

$$I_0(h) = \sum_{i=0}^{\infty} \frac{h^{2i}}{2^i (i!)^2}, \tag{4.2}$$

the modified Bessel function.

Theorem 4.1 (Ray-Knight theorem for continuous-time random walks)**.** *Let ℓ_T defined in* (1.1) *be the local times of continuous-time simple random walk $(X_t)_{t>0}$ on $\mathbb{Z}$. Let $b \in \mathbb{N}$ and $h > 0$.*

(i) *Under $\mathbb{P}_0$, the process $\big(\ell_{T_b^h}(b-x)\big)_{x=0}^b$ is a time-homogeneous discrete-time Markov chain on $(0,\infty)$, starting at h, with transition density given by*

$$f(h_1, h_2) = \mathrm{e}^{-h_1-h_2} I_0\big(2\sqrt{h_1 h_2}\big), \qquad h_1, h_2 \in (0,\infty). \tag{4.3}$$

(ii) *Under $\mathbb{P}_0$, the processes $\big(\ell_{T_b^h}(b+x)\big)_{x\in\mathbb{N}_0}$ and $\big(\ell_{T_b^h}(-x)\big)_{x\in\mathbb{N}_0}$ are time-homogeneous discrete-time Markov chains on $[0,\infty)$ with transition probabilities given by*

$$P^\star(h_1, \mathrm{d}h_2) = \mathrm{e}^{-h_1}\delta_0(\mathrm{d}h_2) + \mathrm{e}^{-h_1-h_2}\sqrt{\frac{h_1}{h_2}} I_0'\big(2\sqrt{h_1 h_2}\big)\,\mathrm{d}h_2, \qquad h_1, h_2 \in [0,\infty). \tag{4.4}$$

(iii) *The three Markov chains in (i) and (ii) are independent.*

We note that Theorem 4.1(ii) and an outline of its proof can be found in [MS87, (3.1-2)]. This proof uses an embedding of the random walk into a Brownian motion and the Brownian Ray-Knight theorem; we expect that Theorem 4.1(i) and (iii) can also be proved along these lines. In Appendix A, using the density formula of Theorem 2.1, we provide a proof of Theorem 4.1 that is independent of the Brownian Ray-Knight theorem. This opens up the possibility of producing a new proof of this theorem, via a diffusion approximation of the Markov chains having the transition densities in (4.3) and (4.4). Furthermore, we emphasize that our proof can also be adapted to continous-time random walks on cycle-free graphs and has some potential to be extended to more general graphs. Theorem 2.1 contains far-ranging generalisations of the Ray-Knight idea, which are to be studied in future.

Acknowledgment. DB would like to thank the Natural Sciences and Engineering Research Council of Canada for supporting his research. The work of RvdH was supported in part by Netherlands Organisation for Scientific Research (NWO). WK would like to thank the German Science Foundation for awarding a Heisenberg grant (realized in 2003/04). This project was initiated during an extensive visit of RvdH to the University of British Columbia, Vancouver, Canada.

References

[Abd03] A. Abdesselam, Grassmann-Berezin calculus and theorems of the matrix-tree type. *Adv. Appl. Math.* **33:1**, 51–70 (2004).

[AC04] R. Alicandro and M. Cicalese, A general integral representation result for continuum limits of discrete energies with superlinear growth. *SIAM J. Math. Anal.* **36**, 1–37 (2004).

[Ber87] F.A. Berezin, *Introduction to Superanalysis*. Mathematical Physics and Applied Mathematics **9**, D. Reidel Publishing, Dordrecht (1987).

[BK01] M. Biskup and W. König, Long-time tails in the parabolic Anderson model with bounded potential. *Ann. Probab.* **29:2**, 636-682 (2001).

[BDT95] E. Bolthausen, J.-D. Deuschel, and Y. Tamura, Laplace approximations for large deviations of nonreversible Markov processes. The nondegenerate case. *Ann. Probab.* **23:1**, 236–267 (1995).

[BFS82] D.C. Brydges, J. Fröhlich, and T. Spencer, The random walk representation of classical spin systems and correlation inequalities. *Commun. Math. Phys.* **83**, 123–150 (1982).

[BI03a] D.C. Brydges and J.Z. Imbrie, Branched polymers and dimensional reduction. *Annals of Mathematics* **158**, 1019–1039 (2003).

[BI03b] D.C. Brydges and J.Z. Imbrie, Green's function for a hierarchical self-avoiding walk in four dimensions. *Commun. Math. Phys.* **239:3**, 549–584 (2003).

[BM91] D.C. Brydges and I. Muñoz-Maya, An application of Berezin integration to large deviations. *Jour. Theor. Probab.* **4**, 371–389 (1991).

[DZ98] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd edition. Springer, New York (1998).

[DV75-83] M.D. Donsker und S.R.S. Varadhan, Asymptotic evaluation of certain Markov process expectations for large time, I–IV, *Comm. Pure Appl. Math.* **28**, 1–47, 279–301 (1975), **29**, 389–461 (1979), **36**, 183–212 (1983).

[Dyn83] E.M. Dynkin, Gaussian and non-Gaussian random fields associated with Markov processes. *J. Funct. Anal.* **55**, 344–376 (1983).

[Dyn84a] E.M. Dynkin, Local times and quantum fields. *Seminar on Stochastic Processes*, 1983 (Gainesville, Fla., 1983), 69–83, *Progr. Probab. Statist.*, **7** Birkhäuser Boston, Boston, MA (1984).

[Dyn84b] E.M. Dynkin, Polynomials of the occupation field and related random fields. *J. Funct. Anal.* **58**, 20–52 (1984).

[Eise94] N. Eisenbaum, Dynkin's isomorphism theorem and the Ray-Knight theorems. *Probab. Theory Relat. Fields* **99**, 321–335 (1994).

[EK93] N. Eisenbaum and H. Kaspi, A necessary and sufficient condition for the Markov properties of the local time process. *Ann. Probab.* **21:3**, 1591–1598 (1993).

[EK96] N. Eisenbaum and H. Kaspi, On the Markov property of local time for Markov processes on general graphs. *Stoch. Proc. Appl.* **64**, 153–172 (1996).

[EKMRS00] N. Eisenbaum, H. Kaspi, M. B. Marcus, J. Rosen, and Z. Shi, A Ray-Knight theorem for symmetric Markov processes. *Ann. Probab.* **28:4**, 1781–1796 (2000).

[GKS05] N. Gantert, W. König and Z. Shi, Annealed deviations for random walk in random scenery. Preprint (2005). To appear in *Annales Inst. H. Poincaré: Prob. Stat.*

[Gär77] J. Gärtner, On large deviations from the invariant measure, *Th. Prob. Appl.* **22**, 24–39 (1977).

[GH99] J. Gärtner and F. den Hollander, Correlation structure of intermittency in the parabolic Anderson model, *Probab. Theory Relat. Fields* **114**, 1–54 (1999).

[GH93] A. Greven and F. den Hollander, A variational characterization of the speed of a one-dimensional self-repellent random walk, *Ann. Appl. Probab.* **3**, 1067-1099 (1993).

[HKM05] R. van der Hofstad, W. König and P. Mörters, The universality classes in the parabolic Anderson model. Preprint (2005). To appear in *Commun. Math. Phys.*

[KS91] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd edition. Springer, New York (1991).

[KKS75] H. Kesten, M.V. Kozlov and F. Spitzer, A limit law for random walk in random environment. *Compositio Math.* **30**, 145–168 (1975).

[Kni63] F.B. Knight, Random walks and a sojourn density process of Brownian motion, *Trans. Amer. Soc.* **109**, 56–86 (1963).

[LeJ87] Y. Le Jan, Temps local et superchamps, *Séminaire des Probabilités XXI* , Lecture Notes in Math., **1247**, Springer, Berlin, 176–190 (1987).

[Lut83] J. M. Luttinger, The asymptotic evaluation of a class of path integrals. II. *J. Math. Phys.*, **24:8**, 2070–2073 (1983).

[MS87] P. March and A.-S. Sznitman, Some connections between excursion theory and the discrete Schrödinger equation with random potentials. *Probab. Theory Relat. Fields* **109**, 11–53 (1987).

[MR96] M. B. Marcus and J. Rosen, Gaussian chaos and sample path properties of additive functionals of symmetric Markov processes. *Ann. Probab.* **24:3**, 1130–1177 (1996).

[McK80] A.J. McKane, Reformulation of $n \to 0$ models using anticommuting scalar fields. *Physics Lett. A* **76**:22 (1980).

[PS80] G. Parisi and N. Sourlas, Self-avoiding walk and supersymmetry. *J. Physique. Lettres.* **41**, L403–L406 (1980).

[Ray63] D. Ray, Sojourn times of diffusion processes, *Illinois J. Math.* **7**, 615-630 (1963).

[RY91] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion.* Springer, Berlin (1991).

[She85] P. Sheppard, On the Ray-Knight Markov property of local times. *J. London Math. Soc.* **31**, 377–384 (1985).

[Sym69] K. Symanzik, Euclidean quantum theory. In R. Jost, editor, *Local Quantum Theory.* Academic Press, New York, London (1969).

[Toth96] B. Tóth, Generalized Ray-Knight theory and limit theorems for self-interacting random walks on $\mathbb{Z}^1$. *Ann. Probab.* **24:3**, 1324–1367 (1996).

A. Appendix: Proof of Theorem 4.1

In this section, we prove Theorem 4.1. (Recall the discussion in Section 4.2.) To prove Theorem 4.1, we will need the following proposition, which is of independent interest. Recall that the matrix B is the off-diagonal part of the generator A of the Markov chain. Let

$$g_{x,y}(t) = \int_{[0,2\pi]} \mathrm{e}^{t(B_{x,y}\,\mathrm{e}^{\mathrm{i}\theta} + B_{y,x}\,\mathrm{e}^{-\mathrm{i}\theta})}\,\frac{\mathrm{d}\theta}{2\pi}, \qquad t > 0, x, y \in \mathbb{Z}. \tag{A.1}$$

The following does not assume that we are dealing with simple random walk, but holds for any continuous-time nearest-neighbor Markov chain on $\mathbb{Z}$. That is, the generator A is a *tridiagonal* matrix, which means that $A_{x,y} = 0$ for any x, y satisfying $|x - y| \geq 2$.

Proposition A.1. *Assume that the conservative generator $A = (A_{x,y})_{x,y\in\mathbb{Z}}$ of the Markov chain is a tridiagonal matrix. Let $R \subset \mathbb{Z}$ be a finite interval and let $a, b \in R$ with $a \leq b$. Then*

$$\rho_{ab}^{(R)}(l) = \mathrm{e}^{\sum_{x\in R} A_{x,x} l_x} \Big[\prod_{\substack{x<a\\x\in R}} \partial_{l_x} g_{x,y}(\sqrt{l_x l_y})\Big]\Big[\prod_{a\leq x<b} g_{x,y}(\sqrt{l_x l_y})\Big]\Big[\prod_{\substack{y>b\\y\in R}} \partial_{l_y} g_{x,y}(\sqrt{l_x l_y})\Big], \tag{A.2}$$

where $y = x + 1$.

Proposition A.1 says that, for any nearest-neighbor Markov chain, the distribution of the sequence of local times possesses a product structure and may be divided into the piece between starting and ending point and the two boundary pieces. All the three pieces have Markovian structure, with explicit identification of the transition probability function, which is in general inhomogeneous.

The proof of Proposition A.1 makes use of the following two lemmas. The first one is purely algebraic and shows that the determinant of a tridiagonal matrix, after erasing one row and one column, naturally decomposes in a product of three parts.

Lemma A.2. *Let* $R \subset \mathbb{Z}$ *be a finite interval and let* $a, b \in R$ *with* $a \leq b$. *Let* M *be an* $R \times R$ *tridiagonal matrix. Then*

$$\det_{ab}(M) = \det_{aa}^{(\leq a)}(M) \Big(\prod_{a \leq i < b} M_{i+1,i} \Big) \det_{bb}^{(\geq b)}(M), \tag{A.3}$$

where the superscripts $\leq a$ *and* $\geq b$ *denote the sets* $R^{(\leq a)} = \{i \in R \colon i \leq a\}$ *and* $R^{(\geq b)} = \{i \in R \colon i \geq b\}$ *respectively. If the product is over the empty set or if the cofactors apply to empty matrices, then the corresponding factor is set equal to* 1.

Proof. Let $M^{(<a)} = (M_{i,j})_{i,j<a}$ and likewise for other inequalities as superscripts. Let $\widehat{M}$ denote the matrix obtained by removing row a and column b from M so that $\widehat{M}_{i,j} = M_{f(i),g(j)}$, where $f(i) = i$ for $i < a$ and $f(i) = i + 1$ for $i \geq a$. Likewise $g(j) = j$ for $j < b$ and $g(j) = j + 1$ for $j \geq b$. In the proof, we distinguish two separate cases, depending on whether $a = b$ or $a < b$.

Case $a = b$. Then it is easy to see that $\widehat{M} = M^{(<a)} \oplus M^{(>a)}$ is block-diagonal. Since $\det(M^{(<a)}) = \det_{aa}^{(\leq a)}(M)$ and $\det(M^{(>a)}) = \det_{bb}^{(\geq a)}(M)$, whereas the middle term on the right of (A.3) is an empty product, the result (A.3) follows immediately.

Case $a < b$. The $R^{(<a)} \times R^{(<a)}$-submatrix of $\widehat{M}$ (the upper left corner) is equal to $M^{(<a)}$, and right of this block there are throughout zeros in $\widehat{M}$, except for the last row (the row indexed by a). Analogously, the $R^{(\geq b)} \times R^{(\geq b)}$-submatrix of $\widehat{M}$ (the lower right corner) is equal to $\widehat{M}^{(\geq b)} = M^{(>b)}$, and there are throughout zeros above that block in $\widehat{M}$, except for the first column (the column indexed by b). The intermediate block $(\widehat{M}_{i,j})_{a \leq i,j < b}$ is in upper triangular form since, for $i = a, \dots, b-1$ and $j > i$, we have $\widehat{M}_{i,j} = M_{f(i),g(j)} = M_{i+1,j} = 0$. Furthermore, the only non-zero entries below the diagonal of $\widehat{M}$ are in the diagonal that is next to the main diagonal.

We are going to calculate the determinant of $\widehat{M}$ by applying linear row and column operations that transform $\widehat{M}$ into upper triangular form. For $i = \min R, \dots, a-2$, we add a suitable multiple of the i-th row to the $(i+1)$-st row in order that the upper left corner is turned into a upper triangular matrix. Note that these operations do not affect any entry outside this corner. Furthermore, for $j = \max R, \max R - 1, \dots, b+1$, we add a suitable multiple of the j-th column to the $(j-1)$-st column in order that the lower right corner is turned into an upper triangular matrix. Note that these operations do not affect any entry outside this corner.

The resulting $R \times R$-matrix is in upper triangular form, and, since row additions do not change the determinant, its determinant is equal to $\det(\widehat{M}) = \det_{ab}(M)$. This determinant is equal to the product of the three determinants of the left upper corner (which is $\det(M^{(<a)}) = \det_{aa}^{(\leq a)}(M)$), the right lower corner (which is $\det(M^{(>b)}) = \det_{bb}^{(\geq b)}(M)$) and the product of the diagonal entries of the piece inbetween, which is $\prod_{i=a}^{b-1} \widehat{M}_{i,i} = \prod_{a \leq i < b} M_{i+1,i}$. This ends the proof. □

Now we state and prove the second lemma that will be used in the proof of Proposition A.1.

Lemma A.3. *Let V be a finite interval in $\mathbb{Z}$ with $|V|\geq 2$ and let $c\in V$ be the smallest or the largest state. Then, for every $l\in\mathbb{R}_+^V$,*

$$\det\nolimits_{cc}^{(V)}\big(-B+\partial_l\big)\Big[\prod_{x,y\in V\setminus\{c\}} g_{x,y}(\sqrt{l_x l_y})\Big] = 0, \tag{A.4}$$

$$\det\nolimits_{cc}^{(V)}\big(-B+\partial_l\big)\Big[\prod_{x,y\in V} g_{x,y}(\sqrt{l_x l_y})\Big] = \prod_{x,y\in V}\partial_{l_z} g_{x,y}(\sqrt{l_x l_y}), \tag{A.5}$$

where $y=x+1$ and where $z=x$ if c is the largest state and $z=y$ if c is the smallest state in V.

Proof. We first prove (A.4). In Theorem 2.1, choose $F(l)=f(l_c)$ to depend only on l_c and integrate over T, to obtain that

$$\int_0^\infty \mathrm{d}T\;\mathbb{E}_c\big[f(\ell_T(c))1\!\!1_{\{X_T=c\}}1\!\!1_{\{R_T=V\}}\big]=\int \mathrm{d}l_c\; f(l_c)\int \rho_{cc}^{(V)}(l)\,\mathrm{d}^{V\setminus\{c\}}l.$$

In the left hand side the indicator functions of the events $\{R_T=V\}$ and $\{X_0=X_T=c\}$ (recall that $|V|\geq 2$) require the Markov chain to visit c at least two times. Conditioning on the number of visits to c to be n_c, the amount of time $\ell_T(c)$ spent at c is the sum of $n_c\geq 2$ exponential random variables with parameter $A_{c,c}$. The density at 0 of a sum of two or more exponential random variables is zero. Therefore, for $l_c=0$,

$$\int \rho_{cc}^{(V)}(l)\,\mathrm{d}^{V\setminus\{c\}}l=0.$$

It can be seen from (2.33) that $l_c\mapsto\rho_{cc}^{(V)}(l)$ is continuous at $l_c=0$. Therefore, $\rho_{cc}^{(V)}(l)=0$ at $l_c=0$. Also by (2.33), for $l_c=0$,

$$0=\det\nolimits_{cc}^{(V)}\big(-B+\partial_l\big)\Big[\prod_{x,y\in V} g_{x,y}(\sqrt{l_x l_y})\Big].$$

In $\det_{cc}^{(V)}\big(-B+\partial_l\big)$ no derivative with respect to l_c appears. Also $g_{x,y}(0)=1$. Therefore, for $l_c=0$, we may replace V by $V\setminus\{c\}$ under the product sign. But this is (A.4).

Now we prove (A.5). The proof is by induction on the number of elements, n, in V. For $n=2$ the statement is easily checked. This initialises the induction, and we are left to advance it. We give the argument only for the case where c is the smallest vertex in V; the other case is essentially the same. We split the product into

$$\prod_{x,y\in V} g_{x,y}(\sqrt{l_x l_y})=g_{c,c+1}(\sqrt{l_c l_{c+1}})\prod_{x,y\in V\setminus\{c\}} g_{x,y}(\sqrt{l_x l_y})$$

apply the differential operator $\det_{cc}^{(V)}(-B+\partial_l)$ and use the product rule of differentiation for the derivative with respect to $\partial_{l_{c+1}}$. Since $\det_{cc}^{(V)}=\det^{(V\setminus\{c\})}$ is linear in the first row we therefore have

$$\begin{aligned}\text{l.h.s. of (A.5)} &= \big(\partial_{l_{c+1}}g_{c,c+1}(\sqrt{l_c l_{c+1}})\big)\Big(\det\nolimits_{c+1\,c+1}^{(V\setminus\{c\})}\big(-B+\partial_l\big)\prod_{x,y\in V\setminus\{c\}} g_{x,y}(\sqrt{l_x l_y})\Big)\\ &\quad+g_{c,c+1}(\sqrt{l_c l_{c+1}})\Big(\det\nolimits_{cc}^{(V)}\big(-B+\partial_l\big)\prod_{x,y\in V\setminus\{c\}} g_{x,y}(\sqrt{l_x l_y})\Big).\end{aligned}$$

By the induction hypothesis, the first term is equal to

$$\big(\partial_{l_{c+1}}g_{c,c+1}(\sqrt{l_c l_{c+1}})\big)\prod_{x,y\in V\setminus\{c\}}\partial_{l_y}g_{x,y}(\sqrt{l_x l_y})=\prod_{x,y\in V}\partial_{l_y}g_{x,y}(\sqrt{l_x l_y}). \tag{A.6}$$

On the other hand, by (A.4), the second term is equal to zero. This advances the induction, and, thus, completes the proof of (A.5). □

Proof of Proposition A.1. Compare (A.1) and (2.33) to see that

$$\rho_{ab}^{(R)}(l) = \mathrm{e}^{\sum_{x\in R} A_{x,x} l_x} \det\nolimits_{ab}^{(R)}\big(-B+\partial_l\big) \prod_{x,y\in R} g_{x,y}(\sqrt{l_x l_y}), \tag{A.7}$$

where we used our convention $y = x+1$. By Lemma A.2 we rewrite this as

$$\begin{aligned}\rho_{ab}^{(R)}(l) = \mathrm{e}^{\sum_{x\in R} A_{x,x} l_x}\Big[\det\nolimits_{aa}^{(\le a)}\big(-B+\partial_l\big)\prod_{x<a} g_{x,y}(\sqrt{l_x l_y})\Big]\\ \times\Big[\prod_{a\le x<b} g_{x,y}(\sqrt{l_x l_y})\Big]\Big[\det\nolimits_{bb}^{(\ge b)}\big(-B+\partial_l\big)\prod_{y>b} g_{x,y}(\sqrt{l_x l_y})\Big].\end{aligned} \tag{A.8}$$

By Lemma A.3, this equals the right-hand side of (A.2). $\square$

Proof of Theorem 4.1. We specialize Proposition A.1 to simple random walk, whose generator $A = (A_{x,y})_{x,y\in\mathbb{Z}}$ is the Laplace operator, i.e., the tridiagonal matrix satisfying $A_{x,y} = 1$ for $|x-y| = 1$ and $A_{x,x} = -2$ for $x \in \mathbb{Z}$. By (A.1) and since $B_{x,y} = B_{y,x} = 1$,

$$g_{x,y}(t) = \int_{[0,2\pi]} \mathrm{e}^{2t\cos\theta}\,\frac{\mathrm{d}\theta}{2\pi} = I_0(2t). \tag{A.9}$$

Fix a finite interval $R \subseteq \mathbb{Z}$ containing a, b. Proposition A.1 gives that, using that $A_{x,x} = -2$ and using the convention $l_{1+\max R} = l_{-1+\min R} = 0$,

$$\begin{aligned}\rho_{ab}^{(R)}(l)\,\mathrm{d}^R l &= \mathrm{e}^{-l_{\min R}}\Big[\prod_{\substack{x<a\\ x\in R}} \mathrm{e}^{-(l_x+l_y)}\partial_{l_x} I_0(2\sqrt{l_x l_y})\,\mathrm{d}l_x\Big]\Big[\prod_{a\le x<b}\mathrm{e}^{-(l_x+l_y)} I_0(2\sqrt{l_x l_y})\,\mathrm{d}l_x\Big]\\ &\qquad\times\Big[\prod_{\substack{y>b\\ y\in R}}\mathrm{e}^{-(l_x+l_y)}\partial_{l_y} I_0(2\sqrt{l_x l_y})\,\mathrm{d}l_y\Big]\mathrm{e}^{-l_{\max R}}\\ &= \Big[\prod_{\substack{x\le a\\ x\in R}} P^\star(l_x, \mathrm{d}l_{x-1})\Big]\Big[\prod_{a\le x<b} f(l_x, l_{x+1})\,\mathrm{d}l_x\Big]\Big[\prod_{\substack{x\ge b\\ x\in R}} P^\star(l_x,\mathrm{d}l_{x+1})\Big].\end{aligned} \tag{A.10}$$

This is the fixed-time equivalent of Theorem 4.1. To go to the stopping time T_b^h, we claim that, for any measurable set $C \subseteq (0,\infty)^{R\setminus\{b\}}$,

$$\begin{aligned}&\mathbb{P}\big((\ell_{T_b^h}(x))_{x\in R\setminus\{b\}} \in C, \mathrm{supp}(\ell_{T_b^h}) = R, T_b^h \in \mathrm{d}T\big)\,\mathrm{d}h\\ &\quad= \mathbb{P}\big((\ell_T(x))_{x\in R\setminus\{b\}} \in C, \mathrm{supp}(\ell_T) = R, X_T = b, \ell_T(b) \in \mathrm{d}h\big)\,\mathrm{d}T.\end{aligned} \tag{A.11}$$

Indeed, (A.11) implies, for any bounded measurable functional $F\colon (0,\infty)^R \to \mathbb{R}$ and any bounded measurable function $f\colon (0,\infty)\to\mathbb{R}$, that

$$\int_0^\infty f(h)\mathbb{E}_a\big[F(\ell_{T_b^h})\mathbb{1}_{\{R_{T_b^h}=R\}}\big]\,\mathrm{d}h = \int_0^\infty \mathbb{E}_a\big[F(\ell_T)\mathbb{1}_{\{R_T=R\}}\mathbb{1}_{\{X_T=b\}} f(\ell_T(b))\big]\,\mathrm{d}T. \tag{A.12}$$

By Theorem 2.1 and (2.3), the right hand side is equal to $\int_{(0,\infty)^R} F(l) f(l_b)\rho_{ab}^{(R)}(l)\,\mathrm{d}^R l$. Equation (A.10) gives that, on the event $\{\mathrm{supp}(\ell_{T_b^h}) = R\}$, the distribution of $\ell_{T_b^h}$ is the one that is claimed in Theorem 4.1. Hence, Theorem 4.1 follows from (A.11).

We now prove (A.11). It is sufficient to prove that, for every $h > 0$,

$$\begin{aligned}&\mathbb{P}\Big((\ell_{T_b^h}(x))_{x\in R\setminus\{b\}} \in C, \mathrm{supp}(\ell_{T_b^h}) = R, T_b^h \in (T, T+\varepsilon)\Big)\\ &\quad= \mathbb{P}\Big((\ell_T(x))_{x\in R\setminus\{b\}} \in C, \mathrm{supp}(\ell_T) = R, X_T = b, \ell_T(b) \in (h-\varepsilon, h)\Big) + o(\varepsilon), \qquad \text{as } \varepsilon\downarrow 0.\end{aligned} \tag{A.13}$$

Equation (A.13) is a consequence of

$$\begin{aligned}&\mathbb{P}((\ell_{T_b^h}(x))_{x\in R\setminus\{b\}}\in C,\mathrm{supp}(\ell_{T_b^h})=R, T_b^h\in(T,T+\varepsilon))\\ &\quad=\mathbb{P}\big((\ell_{T_b^h}(x))_{x\in R\setminus\{b\}}\in C,\mathrm{supp}(\ell_{T_b^h})=R, X_t=b\,\forall t\in(T,T+\varepsilon),\ell_T(b)\in(h-\varepsilon,h)\big)+o(\varepsilon).\end{aligned}\tag{A.14}$$

Indeed, (A.13) follows from (A.14) since, conditionally on the event $\{T_b^h\in(T,T+\varepsilon)\}$, with high probability, the random walker spends all the time in the interval $(T,T+\varepsilon)$ in b and therefore does not change the local times in any other point than b during that time.

To prove (A.14), note that, on $\{T_b^h\in(T,T+\varepsilon)\}$, we have $\ell_T(b)\leq h<\ell_{T+\varepsilon}(b)$. Since also $\ell_{t+s}(b)\leq\ell_t(b)+s$ for any $t,s\geq 0$, we also have that $\ell_T(b)\in(h-\varepsilon,h]$. This shows that $\{T_b^h\in(T,T+\varepsilon)\}\subseteq\{\ell_T(b)\in(h-\varepsilon,h]\}$. Furthermore, on $\{T_b^h\in(T,T+\varepsilon)\}$, there is a $t\in(T,T+\varepsilon)$ such that $X_t=b$. On this event, the probability of the event $\{\exists s\in(T,T+\varepsilon)\colon X_s\neq b\}$ is $\leq\mathcal{O}(\varepsilon^2)$, since $\ell_T(b)\in(h-\varepsilon,h]$ on this event and at least one step happens during the time interval $(T,T+\varepsilon)$. Since $\ell_T(b)$ has a density, this first event has probability $\leq\mathcal{O}(\varepsilon)$ the second has probability $\leq\mathcal{O}(\varepsilon)$, and by the Markov property at time T, the intersection of these events has probability $\leq\mathcal{O}(\varepsilon^2)$. This ends the proof of (A.14). $\square$