The parabolic Anderson model on Galton-Watson trees

Wang, D.

Appendix of Part I

APPENDIX A

Appendix: Part I

§A.1 Large deviation principle for the local times of Markov renewal processes

The following LDP, which was used in the proof of Lemma 2.3.7, was derived in [31, Proposition 1.2], and generalises the LDP for the empirical distribution of a Markov proceses on a finite state space derived in [15]. See [11, Chapter III] for the definition of the LDP.

Proposition A.1.1. *Let $Y = (Y_t)_{t\geq 0}$ be the Markov renewal process on the finite graph $G = (V, E)$ with transition kernel $(\pi_{x,y})_{\{x,y\}\in E}$ and with sojourn times whose distributions $(\psi_x)_{x\in V}$ have support $(0, \infty)$. For $t > 0$, let L^Y_t denote the empirical distribution of Y at time t (see (2.12)). Then the family $(\mathbb{P}(L^Y_t \in \cdot))_{t>0}$ satisfies the LDP on $\mathcal{P}(V)$ with rate t and with rate function $I^\dagger_E$ given by*

$$I^\dagger_E(p) = \inf_{\beta\in(0,\infty)} \inf_{q\in\mathcal{P}(V)} \left[\widehat{K}(\beta q) + \widetilde{K}(p \mid \beta q)\right] \tag{A.1}$$

with

$$\begin{aligned} \widehat{K}(\beta q) &= \sup_{\widehat{q}\in\mathcal{P}(V)} \sum_{x\in V} \beta q(x) \log\left(\frac{\widehat{q}(x)}{\sum_{y\in V} \pi_{x,y}\widehat{q}(y)}\right), & \text{(A.2)}\\ \widetilde{K}(p \mid \beta q) &= \sum_{x\in V} \beta q(x)\, (\mathcal{L}\lambda_x)\left(\tfrac{p(x)}{\beta q(x)}\right), & \text{(A.3)}\end{aligned}$$

where

$$\begin{aligned} (\mathcal{L}\lambda_x)(\alpha) &= \sup_{\theta\in\mathbb{R}}[\alpha\theta - \lambda_x(\theta)], \quad \alpha \in [0, \infty),\\ \lambda_x(\theta) &= \log \textstyle\int_0^\infty \mathrm{e}^{\theta\tau}\psi_x(\mathrm{d}\tau), \qquad \theta \in \mathbb{R}. \end{aligned} \tag{A.4}$$

The rate function $I^\dagger_E$ consist of two parts: $\widehat{K}$ in (A.2) is the rate function of the LDP on $\mathcal{P}(V)$ for the empirical distribution of the discrete-time Markov chain on V with transition kernel $(\pi_{x,y})_{\{x,y\}\in E}$ (see [11, Theorem IV.7]), while $\widetilde{K}$ in (A.3) is the rate function of the LDP on $\mathcal{P}(0, \infty)$ for the empirical mean of the sojourn times, given the empirical distribution of the discrete-time Markov chain. Moreover, λ_x is the cumulant generating function associated with ψ_x, and $\mathcal{L}\lambda_x$ is the Legendre transform of λ_x, playing the role of the Cramèr rate function for the empirical mean of the i.i.d.

sojourn times at x. The parameter β plays the role of the ratio between the continuous time scale and the discrete time scale.

§A.2 Sojourn times: cumulant generating functions and Legendre tranforms

In Appendix A.2.1 we recall general properties of cumulant generating functions and Legendre transforms, in Appendices A.2.2 and A.2.3 we identify both for the two sojourn time distributions arising in Lemma 2.3.7, respectively.

§A.2.1 General observations

Let λ be the cumulant generating function of a non-degenerate sojourn time distribution ϕ, and $\mathcal{L}\lambda$ be the Legendre transform of λ (recall (2.25)). Both λ and $\mathcal{L}\lambda$ are strictly convex, are analytic in the interior of their domain, and achieve a unique zero at $\theta = 0$, respectively, $\alpha = \alpha_c$ with $\alpha_c = \int_0^\infty \tau\phi(\mathrm{d}\tau)$. Furthermore, λ diverges at some $\theta_c \in (0, \infty]$ and has slope α_c at $\theta = 0$. Moreover, if the slope of λ diverges at θ_c, then $\mathcal{L}\lambda$ is finite on $(0, \infty)$.

The supremum in the Legendre transform defining $(\mathcal{L}\lambda)(\alpha)$ is uniquely taken at $\theta = \theta(\alpha)$ solving the equation

$$\lambda'(\theta(\alpha)) = \alpha, \quad \alpha > 0.$$

The tangent of λ with slope α at $\theta(\alpha)$ intersects the vertical axis at $(-\mathcal{L}\lambda)(\alpha)$, i.e., putting

$$\mu(\alpha) = \lambda(\theta(\alpha)) \tag{A.5}$$

we have

$$\mu(\alpha) = \alpha(\mathcal{L}\lambda)'(\alpha) - (\mathcal{L}\lambda)(\alpha). \tag{A.6}$$

(See Fig. A.1.) Note that by differentiating (A.6) we get

$$\mu'(\alpha) = \alpha(\mathcal{L}\lambda)''(\alpha),$$

which shows that $\alpha \mapsto \mu(\alpha)$ is strictly increasing and hence invertible, with inverse function μ^{-1}. Note that by differentiating the relation $(\mathcal{L}\lambda)(\alpha) = \alpha\theta(\alpha) - \lambda(\theta(\alpha))$ we get

$$(\mathcal{L}\lambda)'(\alpha) = \theta(\alpha). \tag{A.7}$$

A further relation that is useful reads

$$(\mathcal{L}\lambda)' \circ \mu^{-1} = \lambda^{-1}, \tag{A.8}$$

which follows because $\mu = \lambda \circ \theta$ by (A.5) and $(\mathcal{L}\lambda)' = \theta$ by (A.7).

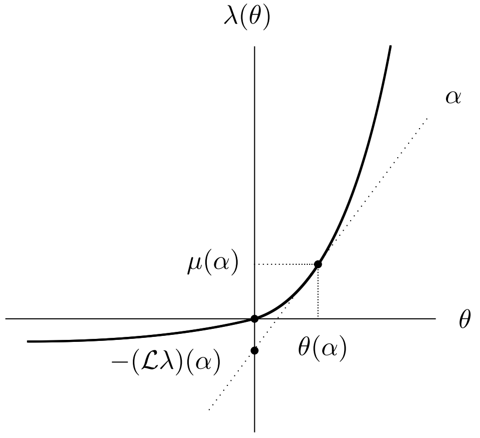

Figure A.1: Picture exhibiting the link between $\lambda(\theta)$, $(\mathcal{L}\lambda)(\alpha)$, $\theta(\alpha)$, $\mu(\alpha)$. The dotted line is the tangent of λ with slope α, crossing the horizontal axis at $-(\mathcal{L}\lambda)(\alpha)$, and touching λ at the point $(\theta(\alpha), \mu(\alpha))$. All are analytic on the interior of their domain.

§A.2.2 Exponential sojourn time

If $\phi = \mathrm{EXP}(d+1)$, then the cumulant generating function $\lambda(\theta) = \log \int_0^\infty \mathrm{e}^{\theta\tau}\psi(\mathrm{d}\tau)$ is given by

$$\lambda(\theta) = \begin{cases} \log\left(\frac{d+1}{d+1-\theta}\right), & \theta < d+1, \\ \infty, & \theta \geq d+1. \end{cases}$$

To find $(\mathcal{L}\lambda)(\alpha)$, we compute

$$\frac{\partial}{\partial\theta}[\alpha\theta - \log(\tfrac{d+1}{d+1-\theta})] = \alpha - \frac{1}{d+1-\theta}, \qquad \frac{\partial^2}{\partial\theta^2}[\alpha\theta - \log(\tfrac{d}{d+1-\theta})] = -\frac{1}{(d+1-\theta)^2} < 0.$$

Hence the supremum in (2.24) is uniquely taken at

$$\theta(\alpha) = d+1-\tfrac{1}{\alpha}, \qquad \alpha > 0,$$

so that

$$(\mathcal{L}\lambda)(\alpha) = \alpha(d+1) - 1 - \log[\alpha(d+1)], \qquad \alpha > 0. \tag{A.9}$$

Thus, λ and $\mathcal{L}\lambda$ have the shape in Fig. A.2, with $\theta_c = d+1$ and $\alpha_c = \frac{1}{d+1}$, and with $\lim_{\theta\uparrow\theta_c}\lambda(\theta) = \infty$ and $\lim_{\theta\uparrow\theta_c}\lambda'(\theta) = \infty$.
Note that μ has domain $(0,\infty)$ and range $\mathbb{R}$.

§A.2.3 Non-exponential sojourn time

For $\phi = \psi$ the computations are more involved. Let $\mathcal{T}^* = (E, V)$ be the infinite rooted regular tree of degree $d+1$. Write $\mathcal{O}$ for the root. Let $X = (X_n)_{n\in\mathbb{N}_0}$ be the discrete-time simple random walk on $\mathcal{T}^* = (E, V)$ starting from $\mathcal{O}$. Write $\tau_{\mathcal{O}}$ to denote the

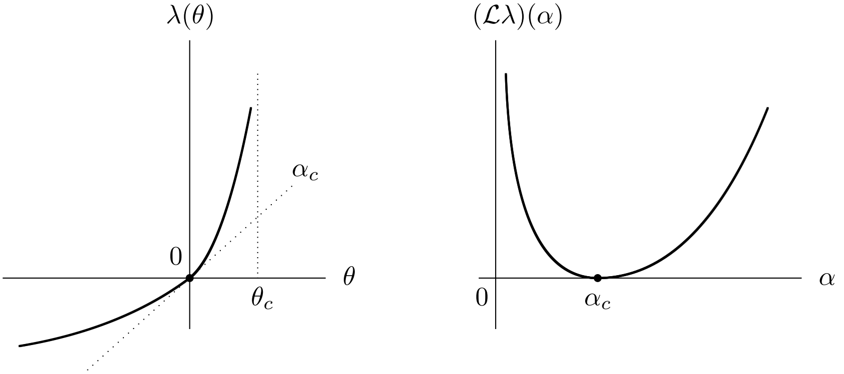

Figure A.2: Picture of $\theta \mapsto \lambda(\theta)$ (left) and $\alpha \mapsto (\mathcal{L}\lambda)(\alpha)$ (right) for $\phi = \mathrm{EXP}(d+1)$.

time of the *first return* of X to $\mathcal{O}$. Define $r = \mathbb{P}_{\mathcal{O}}(\tau_{\mathcal{O}} < \infty)$. It is easy to compute r by projecting X on $\mathbb{N}_0$: r is the return probability to the origin of the random walk on $\mathbb{N}_0$ that jumps to the right with probability $p = \frac{d}{d+1}$ and to the left with probability $q = \frac{1}{d+1}$, which equals $\frac{p}{q}$ (see [33, Section 8]). Thus, $r = \frac{1}{d}$.

For $y \in \mathcal{T}^*$, define $h_y = \mathbb{P}_y(\tau_{\mathcal{O}} < \infty)$. Then h_y can be explicitly calculated, namely,

$$h_y = \begin{cases} d^{-|y|}, & y \in \mathcal{T}^* \setminus \{\mathcal{O}\}, \\ 1, & y = \mathcal{O}. \end{cases}$$

Note that h is a harmonic function on $\mathcal{T}^* \setminus \mathcal{O}$, i.e., $h_y = \sum_{z \in \mathcal{T}^*} \widehat{\pi}_{y,z} h_z$, $y \in \mathcal{T}^* \setminus \mathcal{O}$. We can therefore consider the Doob-transform of X, which is the random walk with transition probabilities away from the root given by

$$\check{\sigma}_{y,z} = \begin{cases} \frac{d}{d+1}, & z = y^{\uparrow}, \\ \frac{1}{d}\frac{1}{d+1}, & z \neq y^{\uparrow}, \{y,z\} \in E, \\ 0, & \text{else}, \end{cases} \qquad y \in \mathcal{T}^* \setminus \{\mathcal{O}\},$$

and transition probabilities from the root are given by

$$\check{\sigma}_{\mathcal{O},z} = \begin{cases} \frac{1}{d}, & \{\mathcal{O}, z\} \in E, \\ 0, & \text{else}. \end{cases}$$

Thus, the Doob-transform reverses the upward and the downward drift of X.

Recall from Lemma 2.3.7 that ψ is the distribution of $\tau_{\mathcal{O}}$ *conditional* on $\{\tau_{\mathcal{O}} < \infty\}$ and on X leaving $\mathcal{O}$ at time 0.

Lemma A.2.1. *Let* $\lambda(\theta) = \log \int_0^\infty \mathrm{e}^{\theta\tau} \psi(\mathrm{d}\tau)$. *Then*

$$\mathrm{e}^{\lambda(\theta)} = \begin{cases} \frac{d+1-\theta}{2} \left[1 - \sqrt{1 - \frac{4d}{(d+1-\theta)^2}}\,\right], & \theta \in (-\infty, \theta_c], \\ \infty, & \textit{else}, \end{cases} \tag{A.10}$$

with $\theta_c = (\sqrt{d}-1)^2$. *The range of* $\exp\circ\lambda$ *is* $(0,\sqrt{d}\,]$, *with the maximal value is uniquely taken at* $\theta=\theta_c$.

Proof. To compute the moment-generating function of $\tau_{\mathcal{O}}$, we consider the Doob-transform of X and its projection onto $\mathbb{N}_0$. Let $p_{2k}=P(\tau_{\mathcal{O}}=2k)$. It is well-known that (see [33, Section 8])

$$G^{p,q}(s)=\mathbb{E}(s^{\tau_{\mathcal{O}}}\mid \tau_{\mathcal{O}}<\infty)=\sum_{k\in\mathbb{N}} s^{2k}p_{2k}=\frac{1}{2p}\left[1-\sqrt{1-4pqs^2}\right], \qquad |s|\le 1. \tag{A.11}$$

Therefore we have

$$\begin{aligned} e^{\lambda(\theta)} &= \mathbb{E}(e^{\theta\tau_{\mathcal{O}}})=\sum_{k\in\mathbb{N}} p_{2k}\left[\mathbb{E}\left(e^{\theta\,\mathrm{EXP}(d+1)}\right)\right]^{2k-1} \\ &= \sum_{k\in\mathbb{N}} p_{2k}\left(\frac{d+1}{d+1-\theta}\right)^{2k-1} = \left(\frac{d+1-\theta}{d+1}\right)G^{p,q}(s) \end{aligned} \tag{A.12}$$

with

$$p=\tfrac{1}{d+1}, \qquad q=\tfrac{d}{d+1}, \qquad s=\frac{d+1}{d+1-\theta}.$$

Inserting (A.11) into (A.12), we get the formula for $\lambda(\theta)$. From the term in the square root we see that $\lambda(\theta)$ is finite if and only if $\theta\le\theta_c=d+1-2\sqrt{d}=(\sqrt{d}-1)^2$. □

There is no easy closed form expression for $(\mathcal{L}\lambda)(\alpha)$, but it is easily checked that λ and $\mathcal{L}\lambda$ have the shape in Fig. A.3, with $\theta_c=(\sqrt{d}-1)^2$ and $\alpha_c=\int_0^\infty \tau\psi(d\tau)<\infty$, and with $\lambda(\theta_c)=\log\sqrt{d}<\infty$ and $\lambda'(\theta_c)=\infty$, i.e., there is a *cusp* at the threshold θ_c, implying that $\mathcal{L}\lambda$ is finite on $(0,\infty)$. It follows from (A.7) that

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}(\mathcal{L}\lambda)(\alpha)=\lim_{\alpha\to\infty}\theta(\alpha)=\theta_c. \tag{A.13}$$

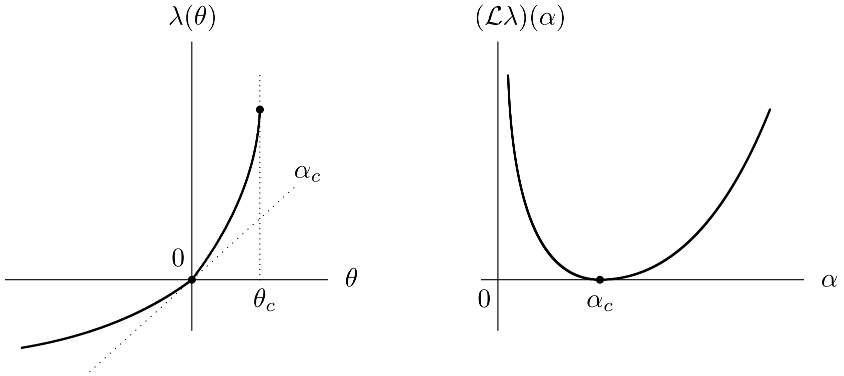

Figure A.3: Picture of $\theta\mapsto\lambda(\theta)$ *(left) and* $\alpha\mapsto(\mathcal{L}\lambda)(\alpha)$ *(right) for* $\phi=\psi$.

APPENDIX A

Lemma A.2.2. *The function* $\lambda^{-1} \circ \log = (\exp \circ \lambda)^{-1}$ *is given by*

$$(\exp \circ \lambda)^{-1}(\beta) = d + 1 - \beta - \frac{d}{\beta}, \qquad \beta \in (0, \sqrt{d}]. \tag{A.14}$$

The range of $(\exp \circ \lambda)^{-1}$ *is* $(-\infty, \theta_c]$*, with the maximal value* θ_c *uniquely taken at* $\beta = \sqrt{d}$*.*

Proof. We need to invert $\exp \circ \lambda$ in (A.10). Abbreviate $\chi = \frac{d+1-\theta}{2}$. Then

$$\beta = \chi \left[1 - \sqrt{1 - \frac{d}{\chi^2}} \right] \quad \Longrightarrow \quad \chi = \frac{\beta^2 + d}{2\beta} \quad \Longrightarrow \quad \theta = d + 1 - \frac{\beta^2 + d}{\beta}.$$

□

Note that $(\sqrt{d}, \infty)$ is not part of the domain of $(\exp \circ \lambda)^{-1}$, even though the right-hand side of (A.14) still makes sense (as a second branch). Note that μ has domain $(0, \infty)$ and range $(-\infty, \sqrt{d}]$ (see Fig. A.1).

§A.3 Large deviation estimate for the local time away from the backbone

In this appendix we derive a large deviation principle for the *total local times at successive depths* of the random walk on $\mathcal{T}^{\mathbb{Z}}$ (see Fig. 2.3). This large deviation principle is not actually needed, but serves as a warm up for the more elaborate computations in Section 2.4.

For $k \in \mathbb{N}_0$, let V_k be the set of vertices in $\mathcal{T}^{\mathbb{Z}}$ that are at distance k from the backbone (see Fig. 2.3). For $R \in \mathbb{N}$, define

$$\begin{aligned} \ell_t^R(k) &= \textstyle\sum_{x \in V_k} \ell_t^{\mathbb{Z}}(x), & k = 0, 1, \ldots, R, \\ \ell_t^R &= \textstyle\sum_{k > R} \sum_{x \in V_k} \ell_t^{\mathbb{Z}}(x), & k = R + 1, \end{aligned}$$

and

$$L_t^R = \frac{1}{t} \Big((\ell_t(k))_{k=0}^R, \ell_t^R \Big).$$

Abbreviate $V_R^* = \{0, 1, \ldots, R, R+1\}$,

Lemma A.3.1. *For every* $R \in \mathbb{N}$, $(L_t^R)_{t \geq 0}$ *satisfies the large deviation principle on* $\mathcal{P}(V_R^*)$ *with rate* t *and with rate function* $I_R^\dagger$ *given by*

$$\begin{aligned} I_R^\dagger(p) = \big[\sqrt{(d-1)p(0)} - \sqrt{dp(1)} \big]^2 + \sum_{k=1}^{R-1} \big[\sqrt{p(k)} - \sqrt{dp(k+1)} \big]^2 \\ + \big[\sqrt{p(R) + p(R+1)} - \sqrt{dp(R+1)} \big]^2. \end{aligned} \tag{A.15}$$

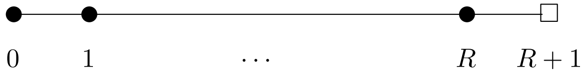

Figure A.4: Depths $k = 0, 1, \dots, R$ and $k > R$.

Proof. By monitoring the random walk on the tree in Fig. 2.3 and projecting its depth on the vertices $0, 1, \dots, R$, respectively, $R + 1$, we can apply the LDP in Proposition A.1.1 (see Fig. A.4).

1. The sojourn times have distribution $\mathrm{EXP}(d + 1)$ at vertices $k = 0, 1, \dots, R$ and distribution ψ at vertex $k = R + 1$. The transition probabilities are

$$\begin{aligned} &\pi_{0,0} = \tfrac{2}{d+1}, && \pi_{0,1} = \tfrac{d-1}{d+1}, \\ &\pi_{k,k+1} = \tfrac{1}{d+1}, && \pi_{k,k-1} = \tfrac{d}{d+1}, \qquad k = 1, \dots, R, \\ &\pi_{R+1,R} = 1. \end{aligned}$$

Proposition A.1.1 therefore yields that $(L_t^R)_{t \geq 0}$ satisfies the LDP on on $\mathcal{P}(V_R^*)$ with rate t and with rate function $I_R^\dagger$ given by

$$I_R^\dagger(p) = (d+1) \sum_{k=0}^{R} p(k) + \inf_{v\colon V_R^* \to (0,\infty)} \sup_{u\colon V_R^* \to (0,\infty)} L(u,v) \tag{A.16}$$

with

$$L(u,v) = -A - B - C, \tag{A.17}$$

where

$$\begin{aligned} A &= \sum_{k=1}^{R} v(x) \left\{ 1 + \log\left(\frac{du(k-1) + u(k+1)}{u(k)} \, \frac{p(k)}{v(k)} \right) \right\}, \\ B &= v(0) \left\{ 1 + \log\left(\frac{2u(0) + (d-1)u(1)}{u(0)} \, \frac{p(0)}{v(0)} \right) \right\}, \\ C &= v(R+1) \left\{ \log\left(\frac{u(R)}{u(R+1)} \right) - (\mathcal{L}\lambda) \left(\frac{p(R+1)}{v(R+1)} \right) \right\}. \end{aligned}$$

Here we use (A.9) to compute A and B, and for C we recall that $\mathcal{L}\lambda$ is the Legendre transform of the cumulant generation function λ of ψ computed in Lemma A.10.

2. We compute the infimum of $L(u, v)$ over v for fixed u.

- For $k = 1, \dots, R$,

$$\begin{aligned} \frac{\partial A}{\partial v(k)} &= \log\left(\frac{du(k-1) + u(k+1)}{u(k)} \, \frac{p(k)}{v(k)} \right), \\ \Longrightarrow \bar{v}_u(k) &= p(k) \, \frac{du(k-1) + u(k+1)}{u(k)}. \end{aligned}$$

The second derivative is $1/v(k) > 0$.

• For $k = 0$,

$$\frac{\partial B}{\partial v(0)} = \log\left(\frac{2u(0)+(d-1)u(1)}{u(0)}\,\frac{p(0)}{v(0)}\right),$$
$$\Longrightarrow \bar{v}_u(0) = p(0)\,\frac{2u(0)+(d-1)u(1)}{u(0)}.$$

The second derivative is $1/v(0) > 0$.

• For $k = R+1$, the computation is more delicate. Define (recall (A.6) in Appendix A.2)

$$\mu(\alpha) = \alpha(\mathcal{L}\lambda)'(\alpha) - (\mathcal{L}\lambda)(\alpha).$$

The function μ has range $(-\infty, \log\sqrt{d}]$, with the maximal value uniquely taken at $\alpha = \infty$. Therefore there are two cases.

► $u(R+1)/u(R) \le \sqrt{d}$. Compute

$$\frac{\partial C}{\partial v(R+1)} = \mu\left(\frac{p(R+1)}{v(R+1)}\right) - \log\left(\frac{u(R+1)}{u(R)}\right),$$
$$\Longrightarrow \bar{v}(R+1) = \frac{p(R+1)}{\alpha_u(R+1)}$$

with $\alpha_u(R+1)$ solving the equation

$$\log\left(\frac{u(R+1)}{u(R)}\right) = \mu\big(\alpha_u(R+1)\big).$$

Since $\mu'(\alpha) = \alpha(\mathcal{L}\lambda)''(\alpha)$ and $\mathcal{L}\lambda$ is strictly convex (see Fig. A.3 in Appendix A.2), μ is strictly increasing and therefore invertible. Consequently,

$$\alpha_u(R+1) = \mu^{-1}\left(\log\left(\frac{u(R+1)}{u(R)}\right)\right). \tag{A.18}$$

Putting (A.17)–(A.18) together, we get

$$L(u) = \inf_{v\colon V_R^* \to (0,\infty)} L(u,v) = -\sum_{k=1}^{R} A_u(k) - B_u + C_u \tag{A.19}$$

with

$$A_u(k) = \frac{du(k-1)+u(k+1)}{u(k)}\,p(k), \qquad k = 1,\ldots,R,$$
$$B_u = \frac{2u(0)+(d-1)u(1)}{u(0)}\,p(0),$$

and

$$\begin{aligned} C_u &= \frac{p(R+1)}{\alpha_u(R+1)}\left[(\mathcal{L}\lambda)\big(\alpha_u(R+1)\big) - \log\left(\frac{u(R+1)}{u(R)}\right)\right] \\ &= \frac{p(R+1)}{\alpha_u(R+1)}\big[(\mathcal{L}\lambda)\big(\alpha_u(R+1)\big) - \mu\big(\alpha_u(R+1)\big)\big] \\ &= p(R+1)\,(\mathcal{L}\lambda)'\big(\alpha_u(R+1)\big) \\ &= p(R+1)\,((\mathcal{L}\lambda)'\circ\mu^{-1})\left(\log\left(\frac{u(R+1)}{u(R)}\right)\right). \end{aligned}$$

In (A.8) in Appendix A.2 we showed that $(\mathcal{L}\lambda)'\circ\mu^{-1} = \lambda^{-1}$. Moreover, in (A.14) in Appendix A.2 we showed that $(\lambda^{-1}\circ\log) = S$ with

$$S(\beta) = d+1-\beta-\frac{d}{\beta}, \qquad \beta\in(0,\sqrt{d}\,]. \tag{A.20}$$

Since S has domain $(0,\sqrt{d}\,]$, $C_u(R+1)$ is only defined when $u(R+1)/u(R)\le\sqrt{d}$, in which case

$$C_u = p(R+1)\,S\left(\frac{u(R+1)}{u(R)}\right). \tag{A.21}$$

► $u(R{+}1)/u(R)\le\sqrt{d}$. In this case $\frac{\partial C}{\partial v(R+1)}>0$, the infimum is taken at $\bar{v}(R{+}1)=0$, and hence (recall (A.13))

$$C_u = p(R+1)\,(\sqrt{d}-1)^2 = p(R+1)\,S(\sqrt{d}). \tag{A.22}$$

Note that the right-hand side does not depend on u. The expressions in (A.21)–(A.22) can be summarised as

$$C_u = p(R+1)\,S\left(\sqrt{d}\wedge\frac{u(R+1)}{u(R)}\right).$$

3. Next we compute the supremum over u of

$$L(u) = L(u,\bar{v}_u) = -A_u - B_u + C_u. \tag{A.23}$$

with $A_u=\sum_{k=1}^R A_u(k)$. We only write down the derivatives that are non-zero.

• For $k=2,\dots,R-1$,

$$-\frac{\partial A_u}{\partial u(k)} = -p(k+1)\,\frac{d}{u(k+1)} - p(k-1)\,\frac{1}{u(k-1)} + p(k)\,\frac{du(k-1)+u(k+1)}{u(k)^2}.$$

• For $k=1$,

$$\begin{aligned} -\frac{\partial A_u}{\partial u(1)} &= -p(2)\,\frac{d}{u(2)} + p(1)\,\frac{du(0)+u(2)}{u(1)^2}, \\ -\frac{\partial B_u}{\partial u(1)} &= -p(0)\,\frac{d-1}{u(0)}. \end{aligned}$$

- For $k = R$,

$$
\begin{aligned}
-\frac{\partial A_u}{\partial u(R)} &= -p(R-1)\,\frac{1}{u(R-1)} + p(R)\,\frac{du(R-1)+u(R+1)}{u(R)^2}, \\
\frac{\partial C_u}{\partial u(R)} &= p(R+1)\,\left[\frac{u(R+1)}{u(R)^2} - \frac{d}{u(R+1)}\right] 1_{\left\{\frac{u(R+1)}{u(R)} \le \sqrt{d}\right\}}.
\end{aligned}
$$

- For $k = 0$,

$$
\begin{aligned}
-\frac{\partial A_u}{\partial u(0)} &= -p(1)\,\frac{d}{u(1)}, \\
-\frac{\partial B_u}{\partial u(0)} &= p(0)\,\frac{(d-1)u(1)}{u(0)^2}.
\end{aligned}
$$

- For $k = R+1$,

$$
\begin{aligned}
-\frac{\partial A_u}{\partial u(R+1)} &= -p(R)\,\frac{1}{u(R)}, \\
\frac{\partial C_u}{\partial u(R+1)} &= p(R+1)\,\left[-\frac{1}{u(R)} + \frac{du(R)}{u(R+1)^2}\right] 1_{\left\{\frac{u(R+1)}{u(R)} \le \sqrt{d}\right\}}.
\end{aligned}
$$

All the first derivatives of $A_u + B_u + C_u$ are zero when we choose

$$
\begin{aligned}
&\bar{u}(0) = \sqrt{(d-1)p(0)}, \qquad \bar{u}(k) = \sqrt{d^k p(k)}, \quad k = 1, \dots, R, \\
&\bar{u}(R+1) = \sqrt{d^{R+1}\,\frac{p(R)p(R+1)}{p(R)+p(R+1)}}.
\end{aligned}
\tag{A.24}
$$

All the second derivatives are strictly negative, and so $\bar{u}$ is the unique maximiser.

4. Inserting (A.24) into (A.19), we get

$$
\begin{aligned}
L(\bar{u}) = L(\bar{u}, \bar{v}_{\bar{u}}) &= -\sum_{k=2}^{R-1} A_{\bar{u}}(k) - \left[A_{\bar{u}}(1) + B_{\bar{u}}\right] - A_{\bar{u}}(R) + C_{\bar{u}} \\
&= -\sum_{k=2}^{R-1} \sqrt{dp(k)}\,\left[\sqrt{p(k-1)} + \sqrt{p(k+1)}\,\right] \\
&\quad - \left[2\sqrt{d(d-1)p(0)p(1)} + 2p(0) + \sqrt{dp(1)p(2)}\,\right] \\
&\quad - \left[\sqrt{dp(R-1)p(R)} + \sqrt{\frac{p(R)}{p(R)+p(R+1)}}\,\sqrt{dp(R)p(R+1)}\right] \\
&\quad + p(R+1)\,S\left(\sqrt{\frac{dp(R+1)}{p(R)+p(R+1)}}\right).
\end{aligned}
$$

Recalling (A.16), (A.20) and (A.23), and rearranging terms, we find the expression in (A.15). □

Note that $I_R^\dagger$ has a unique zero at p given by

$$p(0) = \tfrac{1}{2}, \qquad p(k) = \tfrac{1}{2}(d-1)d^{-k}, \quad k = 1, \dots, R, \qquad p(R+1) = \tfrac{1}{2}d^{-R}.$$

This shows that the fraction of the local time typically spent a distance k away from the backbone decays exponentially fast in k.