ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 43, Number 6, 2013

ON THE NUMBER OF ISOLATED VERTICES IN A GROWING RANDOM GRAPH

ANATOLII A. PUHALSKII

ABSTRACT. This paper studies the properties of the number of isolated vertices in a random graph where vertices arrive one-by-one at times $1, 2, \dots$. They are connected by edges to the previous vertices independently with the same probability. Assuming that the probability of an edge tends to zero, we establish the asymptotics of large, normal, and moderate deviations for the stochastic process of the number of the isolated vertices considered at times inversely proportional to that probability. In addition, we identify the most likely trajectory for that stochastic process to follow conditioned on the event that at a large time the graph is found with a large number of isolated vertices.

1. A problem formulation and main results. The random graph $G(n,p)$, first studied by Gilbert [**6**], is defined as an undirected graph on n vertices where the vertices are linked by edges independently with probability p. A great deal of attention has been paid to the asymptotic properties of the sparse graph $G(n,c/n)$ as $n \to \infty$, see, e.g., Bollobás [**2**], Janson, Łuczak and Ruciński [**10**] and Kolchin [**12**]. In this paper, we investigate the dynamic of the number of the isolated vertices in a growing version of the random graph $G(n,c_n/n)$ where new vertices are added one-by-one. At time 0 no vertices are present and the ith vertex, where $i = 1, 2, \dots$, arrives at time i. It is connected to each of the $i-1$ previous vertices independently with probability c_n/n. Assuming that $c_n \to c > 0$ as $n \to \infty$, we obtain the asymptotics of normal, moderate and large deviations for the suitably time-scaled and normalised stochastic process of the number of the isolated vertices. In addition, we find the most probable trajectory for the process of the number of the isolated vertices to follow before time n conditioned on the event that the number of the isolated vertices at time n is close to

2010 AMS *Mathematics subject classification.* Primary 60C05, 60G99, Secondary 60F10, 60F17.
Keywords and phrases. Random graphs, stochastic processes, isolated vertices, large deviations, weak convergence.
Received by the editors on February 18, 2011, and in revised form on June 20, 2011.
DOI:10.1216/RMJ-2013-43-6-1941

a multiple of n. As a byproduct, new derivations of the asymptotics of normal and large deviations for the number of the isolated vertices of $G(n, c/n)$ are provided.

Let $V_k^{(n)}$ represent the number of the isolated vertices on the arrival of the kth vertex in the graph with edge probability c_n/n. Let $\alpha_{ik}^{(n)}$, for $k > i$, represent a Bernoulli random variable which equals one if the kth vertex connects with the ith present vertex. The $\alpha_{ik}^{(n)}$ are independent with $\mathbf{P}(\alpha_{ik}^{(n)} = 1) = c_n/n$. As mentioned, $V_0^{(n)} = 0$ and $V_1^{(n)} = 1$, so adopting the conventions that $\prod_{\varnothing} = 1$ and $\sum_{\varnothing} = 0$, we can write that

$$V_k^{(n)} = V_{k-1}^{(n)} + \prod_{i=1}^{k-1}(1 - \alpha_{ik}^{(n)}) - \sum_{i=1}^{k-1} \alpha_{ik}^{(n)} \xi_{i,k-1}^{(n)}, \quad k \in \mathbf{N}, \tag{1.1}$$

where $\xi_{i,k-1}^{(n)}$ represents the indicator random variable of the event that the ith vertex out of $k-1$ present is isolated:

$$\xi_{i,k-1}^{(n)} = \prod_{j=1}^{i-1}(1 - \alpha_{ji}^{(n)}) \prod_{j=i+1}^{k-1} (1 - \alpha_{ij}^{(n)}). \tag{1.2}$$

These equations are used to derive our main results.

Let us introduce the stochastic process $X^{(n)}(t) = V_{\lfloor nt \rfloor}^{(n)}/n$, where $t \in \mathbf{R}_+$ and $\lfloor \cdot \rfloor$ denotes the integer part. Evidently, $0 \le X^{(n)}(t) \le \lfloor nt \rfloor /n$. Observing that the process $(X^{(n)}(t),\ t \in \mathbf{R}_+)$ has a separable range, we may and will consider it as a random element of the (nonseparable) Fréchet space $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ of $\mathbf{R}$-valued rightcontinuous functions with lefthand limits defined on $\mathbf{R}_+$ which is endowed with the compact open topology and Borel σ-algebra.

We say that a $[0, \infty]$-valued function $\mathbf{I}$ on a metric space $\mathbf{M}$ is a large deviation rate function if it is lower compact, i.e., the sets $\{z \in \mathbf{M} : \mathbf{I}(z) \le q\}$ are compact for all $q \in \mathbf{R}_+$, and $\inf_{z \in \mathbf{M}} \mathbf{I}(z) = 0$. Suppose $\{X^{(n)}, n = 1, 2, \dots\}$ is a sequence of random elements of $\mathbf{M}$ endowed with the Borel σ-algebra which have distributions $\mathbf{P}^{(n)}$. Let $a_n \to \infty$ as $n \to \infty$. The sequence $\{X^{(n)}, n = 1, 2, \dots\}$ is said to obey the large deviation principle (LDP, for short) for rate a_n with large deviation rate function $\mathbf{I}$ if

$$\limsup_{n \to \infty} \frac{1}{a_n} \log \mathbf{P}^{(n)}(F) \le -\inf_{z \in F} \mathbf{I}(z)$$

for all closed sets $F \subset \mathbf{M}$ and

$$\liminf_{n\to\infty} \frac{1}{a_n} \log \mathbf{P}^{(n)}(G) \geq -\inf_{z\in G} \mathbf{I}(z)$$

for all open sets $G \subset \mathbf{M}$. Convergence in distribution of the $X^{(n)}$ to an $\mathbf{M}$-valued random element X is defined similarly, see, e.g., Jacod and Shiryaev [**7**].

For an $\mathbf{R}$-valued absolutely continuous function $x(\cdot)$ defined on $\mathbf{R}_+$, we denote by $\dot{x}(\cdot)$ a version of the Radon-Nikodym derivative of the associated signed measure with respect to Lebesgue measure, so that $\dot{x}(t)$ exists for all t, is Borel measurable and is specified uniquely almost everywhere with respect to Lebesgue measure.

Theorem 1.1. *Let $c_n \to c > 0$ as $n \to \infty$. The processes $(X^{(n)}(t), t \in \mathbf{R}_+)$ obey the* LDP *in $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ for rate n with large deviation rate function*

$$\mathbf{I}(x(\cdot)) = \int_0^\infty \sup_{\lambda\in\mathbf{R}} \Big(\lambda \dot{x}(t) - \log\Big((e^\lambda - 1)e^{-ct} + e^{c(e^{-\lambda}-1)x(t)}\Big)\Big)\, dt,$$

provided the function $(x(t), t \in \mathbf{R}_+) \in \mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ is absolutely continuous, $x(0) = 0$, $x(t) \geq 0$ for all $t \in \mathbf{R}_+$, and $\dot{x}(t) \leq 1$ for almost all $t \in \mathbf{R}_+$ with respect to Lebesgue measure. If any of these conditions is not met, then $\mathbf{I}(x(\cdot)) = \infty$. The large deviation rate function attains zero at the only function $(te^{-ct}, t \in \mathbf{R}_+)$. It represents a law-of-large-numbers-limit of the $X^{(n)}(\cdot)$, meaning that, for all $\epsilon > 0$ and $T > 0$,

$$\lim_{n\to\infty} \mathbf{P}\Big(\sup_{t\in[0,T]} |X^{(n)}(t) - te^{-ct}| > \epsilon\Big) = 0.$$

Remark 1.1. The theorem asserts, in particular, that $\mathbf{I}$ is a lower compact function on $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ and on the subspace of continuous functions with the subspace topology.

Corollary 1.1. *Assume the hypotheses of Theorem* 1.1. *Given $T > 0$, the random variables $X^{(n)}(T)$ obey the* LDP *in $[0, T]$ for rate n*

with large deviation rate function

$$\mathbf{I}_T(a) = \frac{cT^2}{2}\,(\alpha-1)^2 + (T-a)\log\alpha + a\Big(\log\frac{a}{T} + cT\alpha\Big),$$

α being the unique nonnegative solution of the equation $(1+\alpha(e^{-cT\alpha}-1))T = a$.

Given $a \in [0,T]$, for arbitrary $\varepsilon > 0$,

$$\lim_{\gamma\to 0}\liminf_{n\to\infty}\mathbf{P}\Big(\sup_{t\in[0,T]}|X^{(n)}(t) - x_{a,T}(t)| \le \varepsilon \,\Big|\, |X^{(n)}(T)-a| \le \gamma\Big) = 1,$$

where $x_{a,T}(t) = (1+\alpha(e^{-ct\alpha}-1))t$.

Remark 1.2. One can see that $\alpha = 1$ if $a = Te^{-cT}$ which corresponds to the law of large numbers so that $\mathbf{I}_T(Te^{-cT}) = 0$. Also, $\alpha < 1$ if $a > Te^{-cT}$, in particular, $\alpha = 0$ if $a = T$, and $\alpha > 1$ if $a < Te^{-cT}$.

Remark 1.3. If we indicate explicitly dependence on c in $\mathbf{I}_T(a)$ by writing $\mathbf{I}_T^{(c)}(a)$, then we have the following scaling property $\mathbf{I}_T^{(c)}(a) = T\mathbf{I}_1^{(cT)}(a/T)$. Similarly, $x^{(c)}_{a,T}(tT) = T\,x^{(cT)}_{a/T,1}(t)$. These properties are consequences of scaling properties of the prelimiting sequences.

Theorem 1.2. *Let $\sqrt{n}(c_n - c) \to \theta \in \mathbf{R}$ as $n \to \infty$, where $c > 0$. Then the processes $(Y^{(n)}(t), t \in \mathbf{R}_+)$, where $Y^{(n)}(t) = \sqrt{n}(X^{(n)}(t) - te^{-ct})$, converge in distribution in $\mathbf{D}_{co}(\mathbf{R}_+,\mathbf{R})$ to the diffusion process $(Y(t), t \in \mathbf{R}_+)$ given by the equation*

$$Y(t) = \int_0^t -2\theta s e^{-cs}\,ds - c\int_0^t Y(s)\,ds + \int_0^t \sqrt{(1+cs)e^{-cs} + (2cs-1)e^{-2cs}}\,dW(s),$$

where $(W(t), t \in \mathbf{R}_+)$ is a standard Wiener process. In particular, for $T > 0$, the $Y^{(n)}(T)$ converge in distribution in $\mathbf{R}$ to the normally distributed random variable $N(-\theta T^2 e^{-cT}, (cT^2 - T)e^{-2cT} + Te^{-cT})$.

Remark 1.4. The process $(Y(t), t \in \mathbf{R}_+)$ can be explicitly written as

$$Y(t) = -\theta t^2 e^{-ct} + e^{-ct} \int_0^t e^{cs} \sqrt{(1+cs)e^{-cs} + (2cs-1)e^{-2cs}}\, dW(s).$$

It is a Gaussian process with $\mathbf{E}Y(t) = -\theta t^2 e^{-ct}$ and $\operatorname{cov}(Y(s), Y(t)) = (c\min(s,t)^2 - \min(s,t))e^{-c(s+t)} + \min(s,t)e^{-c\max(s,t)}$.

Theorem 1.3. *Let* $(\sqrt{n}/b_n)(c_n - c) \to \widehat{\theta} \in \mathbf{R}$ *as* $n \to \infty$, *where* $c > 0$, $b_n \to \infty$, *and* $b_n/\sqrt{n} \to 0$. *Then the processes* $(\widehat{Y}^{(n)}(t), t \in \mathbf{R}_+)$, *where* $\widehat{Y}^{(n)}(t) = (\sqrt{n}/b_n)(X^{(n)}(t) - te^{-ct})$, *obey the* LDP *in* $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ *for rate* b_n^2 *with large deviation rate function*

$$\widehat{\mathbf{I}}(y(\cdot)) = \frac{1}{2}\int_0^\infty \frac{(\dot{y}(t) + cy(t) + 2\widehat{\theta}\, te^{-ct})^2}{(1+ct)e^{-ct} + (2ct-1)e^{-2ct}}\, dt$$

if $y(\cdot)$ *is absolutely continuous with* $y(0) = 0$ *and* $\widehat{\mathbf{I}}(y(\cdot)) = \infty$ *otherwise. In particular, for* $T > 0$ *and* $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}\Big(\sup_{t\in[0,T]} |\widehat{Y}^{(n)}(t) + \widehat{\theta} t^2 e^{-ct}| > \epsilon\Big) = 0.$$

Corollary 1.2. *Under the hypotheses of Theorem* 1.3, *for* $T > 0$, *the* $\widehat{Y}^{(n)}(T)$ *obey the* LDP *in* $\mathbf{R}$ *for rate* b_n^2 *with large deviation rate function*

$$\widehat{\mathbf{I}}_T(z) = \frac{1}{2}\frac{(z + \widehat{\theta} T^2 e^{-cT})^2}{(cT^2 - T)e^{-2cT} + Te^{-cT}}.$$

Given $z \in \mathbf{R}$, *for arbitrary* $\varepsilon > 0$,

$$\lim_{\gamma\to 0}\liminf_{n\to\infty} \mathbf{P}\Big(\sup_{t\in[0,T]} |\widehat{Y}^{(n)}(t) - y_{z,T}(t)| \le \varepsilon \Big| |\widehat{Y}^{(n)}(T) - z| \le \gamma\Big) = 1,$$

where

$$y_{z,T}(t) = -\widehat{\theta} t^2 e^{-ct} + \frac{z + \widehat{\theta} T^2 e^{-cT}}{(cT^2 - T)e^{-cT} + T}\Big((ct^2 - t)e^{-ct} + t\Big).$$

Now we give an overview of the related results and describe our methods of proof. The LDP for the $X^{(n)}(1)$ asserted in Corollary 1.1, for the case where $c_n = c$, has been obtained in O'Connell [**15**] who uses combinatorial arguments and gives a different form of the large deviation rate function. The central limit theorem contained in the statement of Theorem 1.2 is a special case of the results in Barbour, Karoński and Ruciński [**1**] who proved that a central limit theorem holds for the number of the isolated vertices of $G(n, p)$ if and only if $n^2p \to \infty$ and $np - \log n \to -\infty$ as $n \to \infty$. Kordecki [**13**] and Punkla and Chaidee [**18**] give the rates of convergence. Pittel [**16**] establishes a central limit theorem for the number of the isolated trees of size $k = 1, 2, \dots$ considered as a stochastic process in k. It is also known that if $p = (c_n + \log n)/n$, where $c_n \to c > 0$, then the number of the isolated vertices of $G(n, p)$ is distributed asymptotically according to the Poisson law with parameter e^{-c}, see Janson, Łuczak and Ruciński [**10**, page 80].

Janson [**8, 9**] provides a functional central limit theorem, akin to Theorem 1.2, for the numbers of the isolated trees of various sizes in the "graph process" where there are n vertices and the edges appear independently at times distributed uniformly on $[0, n]$. In particular, the limit for the process of the number of the isolated vertices is a continuous path zero-mean Gaussian process $(\widetilde{Y}(t), t \in \mathbf{R}_+)$ with $\operatorname{cov}(\widetilde{Y}(s), \widetilde{Y}(t)) = e^{-s-t}(\min(s,t) - 1) + e^{-\max(s,t)}$. It can also be written as $d\widetilde{Y}(t) = -\widetilde{Y}(t)\,dt + \sqrt{e^{-t} + e^{-2t}}\,dW(t)$. One can see certain similarities with the process Y. For instance, if $c = 1$, then $\operatorname{cov}(Y(s), Y(t)) = \min(s,t)\operatorname{cov}(\widetilde{Y}(s), \widetilde{Y}(t))$.

For the proofs, we apply general results on the LDP and convergence in distribution for semimartingales from Puhalskii [**17**] and Jacod and Shiryaev [**7**], respectively. An application of the results from [**7**] to equation (1.1) is fairly straightforward and enables us to establish Theorem 1.2. The proof of Theorem 1.3 is more technically involved but is, in essence, also a routine exercise on applying the results from Puhalskii [**17**]. Corollary 1.2 is obtained as an application of standard methods of the calculus of variations. Significantly more effort is required to prove Theorem 1.1 and Corollary 1.1. The general approach is to apply the results in Puhalskii [**17**] to the set-up of Theorem 1.1 and "project" to obtain Corollary 1.1. However, checking the requirements turns out to be difficult. The main problem is the presence of boundaries as described next.

An LDP like the one in Theorem 1.1 is not too difficult to establish provided the supremum in the integrand is attained and the associated λ represents a bounded function of t. However, having $x(t)$ in the double exponential complicates things. If $x(t)$ tends to zero and $\dot{x}(t)$ is negative, then the optimizing λ tends to $-\infty$. Another source of difficulties is a linear rather than a superlinear growth of the log function in the supremum as $\lambda \to +\infty$. Because of that, when $\dot{x}(t) = 1$, the supremum is "attained" at $\lambda = +\infty$. A recipe for tackling situations where the supremum in the expression for the large deviation rate function either is not attained or is attained at unbounded functions is to approximate the values of $\mathbf{I}$ at "bad" $x(\cdot)$ with the values at "good" $x(\cdot)$. The implementation of this step is complicated by the fact that the integrand in the expression for $\mathbf{I}$ cannot be found explicitly. The approximation is accomplished in several steps and necessitates a detailed study of the properties of $\mathbf{I}$.

Once a trajectorial LDP has been established, a finite-dimensional LDP such as in Corollary 1.1 follows by the continuous mapping principle. It thus requires solving the variational problem of minimizing $\mathbf{I}$ over $x(\cdot)$ with $x(0) = 0$ and $x(T) = a$. By the scaling properties mentioned in Remark 1.3, it suffices to consider the case where $T = 1$. At first sight, this variational problem is a classical Lagrange problem and one can hope to find an optimal trajectory by solving the Euler-Lagrange equation. However, the Lagrangian (i.e., the integrand) is not of class C^1 up to the boundary of its domain of definition. It is not even a Carathéodory function which seems to be a standing hypothesis in the calculus of variations. Therefore, standard results for the Lagrange problem are not applicable, cf. [**3**, page 30] and [**4**, Section 3.4.2]. In addition, due to the constraint $\dot{x}(t) \leq 1$ this is in effect a problem of optimal control. Furthermore, since $x(t)$ is restricted to being between 0 and t, it is a problem with state space constraints, so an optimal trajectory may lie on the boundary. In fact, when $x(1) = 1$, the optimal trajectory is $x(t) = t$. In settings with state-space constraints the necessary conditions for a trajectory to be optimal such as Pontryagin's maximum principle are quite involved, see, e.g., [**20**]. Our approach to tackling the case where $x(1) < 1$ is to begin with finding an optimal trajectory which belongs to the interior of the set of constraints between times 0 and 1 by using a standard form of Pontryagin's maximum principle, see, e.g., [**3**]. This is possible because on the interior of the set of constraints the Lagrangian is of class C^1. We succeed in finding an explicit solution of the associated Hamiltonian system of equations. After that we show that trajectories that reach the boundary between times 0 and 1 cannot be optimal. The latter step is required because in order to prove the second assertion of Corollary 1.1 it is essential that there be a unique optimal trajectory, cf. [**5**]. In the absence of nice convexity properties, no general tools for ascertaining the uniqueness of an optimizer seem to be available. Therefore, we need to use fairly intricate arguments to show that any trajectory that does not belong to the interior of the constraint set yields a greater value of the objective function.

This paper is organized as follows. Sections 2 and 3 lay a foundation for the proofs of the main results. Section 2 is concerned with the study of the properties of the large deviation rate function $\mathbf{I}$. In Section 3, the above described variational problem is solved. Theorem 1.1 and Corollary 1.1 are proved in Section 4, Theorem 1.2 is proved in Section 5 and Theorem 1.3 and Corollary 1.2 are proved in Section 6.

We conclude this section with a list of notation and conventions adopted in the paper. We define $\mathbf{1}_A(x)$ to equal one if $x \in A$ and to equal zero otherwise, if x is an element of the sample space Ω, it is usually omitted; $\mathcal{B}(\mathbf{R})$ represents the Borel σ-algebra on $\mathbf{R}$. Subscripts are used in order to denote partial derivatives with respect to the variable(s) represented in the subscript. The pieces of notation $f(v-)$ and $f(v+)$ refer to limits on the left and on the right, respectively. For real numbers x and y, $x \vee y = \max(x, y)$ and $x \wedge y = \min(x, y)$, $0 \log 0 = 0$. Infima over the empty set are understood to equal infinity. The abbreviation a.e. refers to Lebesgue measure. It is assumed that all random entities are defined on a common complete probability space $(\Omega, \mathcal{F}, \mathbf{P})$. All stochastic processes have rightcontinuous trajectories with lefthand limits.

2. Technical preliminaries. This section collects the properties of the integrand in the expression for $\mathbf{I}$ needed in the proofs of Theorem 1.1 and Corollary 1.1. Let $c > 0$. We introduce, for $t \in \mathbf{R}_+$, $x \in \mathbf{R}$ and $\lambda \in \mathbf{R}$,

$$H(t, x, \lambda) = \log((e^{\lambda} - 1)e^{-ct} + e^{c(e^{-\lambda}-1)(x \vee 0 \wedge t)}). \tag{2.1}$$

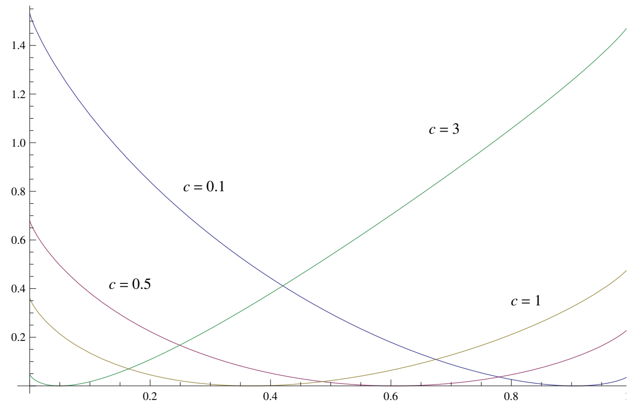

FIGURE 1. The large deviation rate function $I_1(\cdot)$.

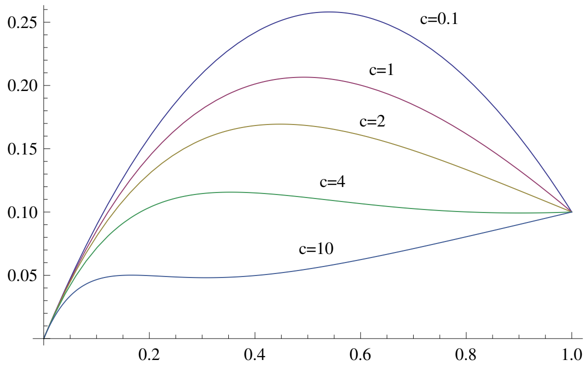

FIGURE 2. Optimal trajectories $x_{a,1}(\cdot)$ for $a = 0.1$.

Since

$$(e^{\lambda}-1)e^{-ct}+e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)}>(1-e^{-ct})\wedge e^{-ct}, \tag{2.2}$$

this function is well defined. It is also infinitely differentiable in λ, the first derivative being given by

$$H_{\lambda}(t,x,\lambda)=\frac{e^{\lambda}e^{-ct}-ce^{-\lambda}(x\vee 0\wedge t)e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)}}{(e^{\lambda}-1)e^{-ct}+e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)}}. \tag{2.3}$$

Lemma 2.1. *The following properties hold.*

1. *The function $H(t,x,\lambda)$ is infinitely differentiable in (t,x,λ) on the domain $\{(t,x,\lambda):0<x<t,\lambda\in\mathbf{R}\}$. It is strictly convex in $\lambda\in\mathbf{R}$ for $(t,x)\in(0,\infty)\times\mathbf{R}$. Moreover, $H_{\lambda\lambda}(t,x,\lambda)>0$ for $(t,x,\lambda)\in(0,\infty)\times\mathbf{R}\times\mathbf{R}$.*

2. *For $t>0$, $x\in\mathbf{R}$ and $\lambda\in\mathbf{R}$, the function $H_{\lambda}(t,x,\lambda)$ is continuous in (t,x,λ), strictly increasing in λ, $H_{\lambda}(t,x,0)=e^{-ct}-c(x\vee 0\wedge t)$, and $\lim_{\lambda\to\infty}H_{\lambda}(t,x,\lambda)=1$. If, in addition, $x>0$, then $\lim_{\lambda\to-\infty}H_{\lambda}(t,x,\lambda)=-\infty$ and, if $x\le 0$, then $\lim_{\lambda\to-\infty}H_{\lambda}(t,x,\lambda)=0$. For $t>0$ and $x\in\mathbf{R}$, the equation $u=H_{\lambda}(t,x,\lambda)$ has at most one solution for λ. The solution exists if and only if either $x>0$ and $u<1$ or $x\le 0$ and $0<u<1$.*

3. *The function $H_{\lambda}(t,x,\lambda)$ is Lipshitz continuous in $x\in\mathbf{R}$ uniformly over $\lambda\ge-\ell$ and $0\le t\le T$ for arbitrary $\ell>0$ and $T>0$.*

Proof. The differentiability properties of $H(t,x,\lambda)$ asserted in part 1 follow by (2.1) and (2.2). Also by (2.1),

$$H(t,x,\lambda)=\log\int_{\mathbf{R}}e^{\lambda y}m(t,x,dy),$$

where, for $t\in\mathbf{R}_+$, $x\in\mathbf{R}$ and $\Gamma\in\mathcal{B}(\mathbf{R})$,

$$\begin{aligned}m(t,x,\Gamma)=e^{-ct}&\mathbf{1}_{\Gamma}(1)\\&+(e^{-c(x\vee 0\wedge t)}-e^{-ct})\mathbf{1}_{\Gamma}(0)\\&+\sum_{k=1}^{\infty}e^{-c(x\vee 0\wedge t)}\,\frac{1}{k!}\,c^k(x\vee 0\wedge t)^k\mathbf{1}_{\Gamma}(-k).\end{aligned}$$

If $t > 0$, then the measure $m(t, x, \cdot)$ is supported by more than point. The Cauchy-Schwartz inequality implies that

$$\int_{\mathbf{R}} e^{\lambda y}\, m(t,x,dy) \int_{\mathbf{R}} y^2 e^{\lambda y}\, m(t,x,dy) > \left(\int_{\mathbf{R}} y e^{\lambda y}\, m(t,x,dy) \right)^2,$$

so $H_{\lambda\lambda}(t, x, \lambda) > 0$ and $H(t, x, \lambda)$ is strictly convex in λ. Part 1 is proved. Part 2 follows by part 1 and (2.3).

To prove the Lipshitz continuity property, it suffices to prove that

$$\sup_{\substack{\lambda \ge -\ell \\ 0<x<t\le T}} |H_{\lambda x}(t,x,\lambda)| < \infty.$$

Calculations show that, for $0 < x < t$ and $\lambda \in \mathbf{R}$,

$$\begin{aligned} H_{\lambda x}(t,x,\lambda) = {} & \frac{ce^{-\lambda}(1-e^{-\lambda})e^{-ct}}{(1-e^{-\lambda})e^{-ct} + e^{-\lambda}e^{c(e^{-\lambda}-1)x}} \\ & - \frac{(e^{-ct} + ce^{-\lambda}x(1-e^{-\lambda})e^{-ct})c(e^{-\lambda}-1)e^{-\lambda}e^{c(e^{-\lambda}-1)x}}{((1-e^{-\lambda})e^{-ct} + e^{-\lambda}e^{c(e^{-\lambda}-1)x})^2} \\ & - ce^{-\lambda}. \end{aligned}$$

Since $(1-e^{-\lambda})e^{-ct} + e^{-\lambda}e^{c(e^{-\lambda}-1)x} \ge e^{-ct}$ for $0 \le x \le t$, we obtain that if $\lambda \ge -\ell$ and $0 < x < t$, then

$$|H_{\lambda x}(t,x,\lambda)| \le ce^{\ell}(1+e^{\ell}) + (1+ce^{\ell}t(1+e^{\ell}))c(e^{\ell}+1)e^{\ell}e^{ce^{\ell}t}e^{ct} + ce^{\ell}. \qquad \square$$

Let, for $t \in \mathbf{R}_+$, $x \in \mathbf{R}$ and $u \in \mathbf{R}$,

$$L(t,x,u) = \sup_{\lambda \in \mathbf{R}} (\lambda u - H(t,x,\lambda)). \tag{2.4}$$

By Lemma 2.1, the maximizer in (2.4), if any, satisfies the equation $u = H_{\lambda}(t, x, \lambda)$. For the solution to exist, it is necessary that $u < 1$. The purpose of the next lemma is to gather some useful inequalities concerning that solution.

Lemma 2.2. *Suppose that $t > 0$ and that λ and u are such that $u = H_{\lambda}(t, x, \lambda)$.*

1. *We have that* $\lambda \geq 0$ (*respectively,* $\lambda \leq 0$) *if and only if* $H_\lambda(t, x, 0) \leq u$ (*respectively,* $H_\lambda(t, x, 0) \geq u$).

2. *If* $\lambda \geq 0$, *then* $\lambda \leq (-\log(1-u)) \vee 0 + \log(1 + c(x \vee 0 \wedge t)) + ct$.

3. *If* $u > 0$, *then*

$$\lambda \geq \left(\frac{1}{2 + c(x \vee 0 \wedge t)}(-\log(1-u) + \log(ct) + ct - \log 2)\right) \\ \wedge (-\log(1-u) + \log u + \log(e^{ct/2} - 1)).$$

4. *If* $\lambda \leq 0$ *and* $x > 0$, *then* $-\lambda \leq \log(e^{-ct} + |u|) - \log(x \vee 0 \wedge t) - \log c$.

5. *If* $\lambda \leq 0$ *and* $u < 0$, *then* $-\lambda \geq \log(-u) - ct$.

6. *If* $u > 0$, *then* $\lambda \geq \log u + \log(e^{ct} - 1) \wedge 0$.

7. *If* $\lambda \leq 0$, $u > 0$ *and* $x > 0$, *then*

$$-\frac{1}{2}(ct + 1) - \frac{1}{2}\log(u + c(x \vee 0 \wedge t)) \\ \leq -\lambda \leq -\frac{1}{2}ct - \frac{1}{2}\log(x \vee 0 \wedge t) - \frac{1}{2}\log c.$$

8. *If* $x \leq 0$ *and* $u > 0$, *then* $\lambda = -\log(1-u) + \log u + \log(e^{ct} - 1)$.

Proof. Part 1 follows by the monotonicity of $H_\lambda(t, x, \lambda)$ in λ (see Lemma 2.1). If $\lambda \geq 0$ and $H_\lambda(t, x, \lambda) \geq 0$, then by (2.3),

$$H_\lambda(t, x, \lambda) \geq \frac{e^\lambda e^{-ct} - c(x \vee 0 \wedge t)}{e^\lambda e^{-ct} + 1}.$$

Thus, if $u \geq 0$ and $\lambda \geq 0$, then $e^\lambda \leq (1 + c(x \vee 0 \wedge t))e^{ct}/(1-u)$. If $\lambda \geq 0$ and $u < 0$, then by (2.3) $e^\lambda e^{-ct} - ce^{-\lambda}(x \vee 0 \wedge t)e^{c(e^{-\lambda}-1)(x \vee 0 \wedge t)} < 0$, so $e^\lambda < c(x \vee 0 \wedge t)e^{ct}$. Part 2 is proved.

An algebraic manipulation of the equation $u = H_\lambda(t, x, \lambda)$ yields, provided $u \geq 0$,

$$\begin{aligned} e^\lambda e^{-ct}(1-u) &= ce^{-\lambda}(x \vee 0 \wedge t)e^{c(e^{-\lambda}-1)(x \vee 0 \wedge t)} \\ &\quad + ue^{c(e^{-\lambda}-1)(x \vee 0 \wedge t)} - ue^{-ct} \\ &\geq ce^{-\lambda}(x \vee 0 \wedge t)e^{c(e^{-\lambda}-1)(x \vee 0 \wedge t)} \\ &\geq ce^{-(1+c(x \vee 0 \wedge t))\lambda}(x \vee 0 \wedge t), \end{aligned} \tag{2.5}$$

so $e^{(2+c(x\vee 0\wedge t))\lambda} \geq ce^{ct}(x\vee 0\wedge t)/(1-u) \geq ce^{ct}t/(2(1-u))$ when $x \geq t/2$. If $x \leq t/2$, then by the equality in (2.5), $e^{\lambda}e^{-ct}(1-u) \geq u(e^{-ct/2}-e^{-ct})$. The assertion of part 3 follows.

If $\lambda \leq 0$, then by (2.3),

$$-|u| \leq H_\lambda(t,x,\lambda) \leq \frac{e^{-ct} - ce^{-\lambda}(x \vee 0 \wedge t)e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)}}{(e^{\lambda}-1)e^{-ct} + e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)}}$$

so

$$-|u|e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)} \leq e^{-ct} - ce^{-\lambda}(x \vee 0 \wedge t)e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)},$$

and $ce^{-\lambda}(x \vee 0 \wedge t) \leq e^{-ct} + |u|$, proving part 4.

By the equality in (2.5),

$$ce^{-\lambda}t \geq e^{\lambda}e^{-ct}(1-u)e^{-c(e^{-\lambda}-1)(x\vee 0\wedge t)} - u + ue^{-ct}e^{-c(e^{-\lambda}-1)(x\vee 0\wedge t)}.$$

If $u \leq 0$ and $\lambda \leq 0$, then the righthand side is greater than $-u + ue^{-ct}$. Hence, the assertion of part 5.

To obtain part 6, we observe that by (2.2) and (2.3),

$$H_\lambda(t,x,\lambda) \leq \frac{e^{\lambda}e^{-ct}}{(1-e^{-ct}) \wedge e^{-ct}}.$$

If $\lambda \leq 0$ and $u \geq 0$, then

$$e^{\lambda}e^{-ct} \geq ce^{-\lambda}(x \vee 0 \wedge t)e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)} \geq ce^{-\lambda}(x \vee 0 \wedge t), \tag{2.6}$$

which implies the righthand inequality of part 7. For the lefthand inequality, we write

$$u \geq \frac{e^{\lambda}e^{-ct}}{e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)}} - ce^{-\lambda}(x \vee 0 \wedge t) \geq e^{\lambda}e^{-ct-1} - ce^{-\lambda}(x \vee 0 \wedge t),$$

where the second inequality follows by (2.6). Part 8 is a straightforward consequence of the definition of λ. □

The next lemma lists the properties of $L(t,x,u)$ pertinent to the proof of Theorem 1.1. If there exists a unique solution to the equation

$u = H_\lambda(t,x,\lambda)$, we let $\widehat{\lambda}(t,x,u)$ represent this solution. We define $\widehat{\lambda}(t,x,u) = 0$ otherwise.

Lemma 2.3. *The function $L(t,x,u)$ is nonnegative. If $t > 0$ and either $x > 0$ and $u < 1$, or $0 < u < 1$, then $L(t,x,u) = \widehat{\lambda}(t,x,u)u - H(t,x,\widehat{\lambda}(t,x,u))$. For $t > 0$, $L(t,x,u)$ is finite if and only if either $x > 0$ and $u \le 1$ or $x \le 0$ and $0 \le u \le 1$. Also, $L(t,x,1) = ct$ for $t \in \mathbf{R}_+$ and $x \in \mathbf{R}$, and $L(t,x,0) = -\log(1-e^{-ct})$ for $t \in \mathbf{R}_+$ and $x \le 0$. The function $\widehat{\lambda}(t,x,u)$ is continuous on $(0,\infty)\times(0,\infty)\times(-\infty,1)$ and on $(0,\infty)\times\mathbf{R}\times(0,1)$, $\widehat{\lambda}(t,x,1-) = \infty$ and $\widehat{\lambda}(t,0,0+) = -\infty$ for $(t,x) \in (0,\infty)\times\mathbf{R}$. The function $L(t,x,u)$ is continuous on $(0,\infty)\times(0,\infty)\times(-\infty,1]$ and on $(0,\infty)\times\mathbf{R}\times[0,1]$.*

Proof. The function $L(t,x,u)$ is nonnegative because $H(t,x,0) = 0$. The representation for $L(t,x,u)$ provided $t > 0$ and either $x > 0$ and $u < 1$, or $0 < u < 1$, follows by Lemma 2.1. If $u > 1$, then $\lim_{\lambda\to\infty}(u - H(t,x,\lambda)/\lambda) = u - 1 > 0$, so $L(t,x,u) = \infty$. If $x \le 0$ and $u < 0$, then $\lim_{\lambda\to-\infty} H(t,x,\lambda) = \log(1-e^{-ct})$ whereas $\lim_{\lambda\to-\infty}\lambda u = \infty$. The equality $L(t,x,1) = ct$ follows because

$$L(t,x,1) = -\inf_{\lambda\in\mathbf{R}} \log(e^{-ct} + e^{-\lambda}(e^{c(e^{-\lambda}-1)(x\vee 0\wedge t)} - e^{-ct})),$$

where the infimum is given by the limit as $\lambda \to \infty$. For $x \le 0$,

$$L(t,x,0) = -\inf_{\lambda\in\mathbf{R}} \log((e^{\lambda}-1)e^{-ct} + 1),$$

where the infimum is given by the limit as $\lambda \to -\infty$.

We establish the continuity properties of $\widehat{\lambda}(t,x,u)$ and $L(t,x,u)$. Suppose that either $(t,x,u) \in (0,\infty)\times(0,\infty)\times(-\infty,1)$ or $(t,x,u) \in (0,\infty)\times\mathbf{R}\times(0,1)$. If $(t_n,x_n,u_n) \to (t,x,u)$ as $n\to\infty$, then by parts 2, 3, and 4 of Lemma 2.2 the sequence $\widehat{\lambda}(t_n,x_n,u_n)$ is bounded, so it has a convergent subsequence whose limit must coincide with $\widehat{\lambda}(t,x,u)$ by the continuity of $H_\lambda(t,x,\lambda)$ and the uniqueness of the solution λ to $u = H_\lambda(t,x,\lambda)$. Hence, $\widehat{\lambda}(t_n,x_n,u_n) \to \widehat{\lambda}(t,x,u)$ as $n\to\infty$ which implies the continuity of $\widehat{\lambda}(t,x,u)$ and of $L(t,x,u)$ on $(0,\infty)\times(0,\infty)\times(-\infty,1)$ and on $(0,\infty)\times\mathbf{R}\times(0,1)$.

To prove the continuity of $L(t,x,u)$ at $(t,x,1)$, where $(t,x)\in(0,\infty)\times\mathbf{R}$, we consider $(t_n,x_n)\to(t,x)$ and $u_n\uparrow 1$ as $n\to\infty$. We have

$$\begin{aligned}L(t_n,x_n,u_n)&=\widehat{\lambda}(t_n,x_n,u_n)(u_n-1)\\&\quad-\log((1-e^{-\hat{\lambda}(t_n,x_n,u_n)})e^{-ct_n}\\&\quad+e^{-\hat{\lambda}(t_n,x_n,u_n)}e^{c(e^{-\hat{\lambda}(t_n,x_n,u_n)}-1)(x_n\vee 0\wedge t_n)}).\end{aligned}$$

By Lemma 2.2 part 3, $\widehat{\lambda}(t_n,x_n,u_n)\to\infty$ as $n\to\infty$. By part 2 of Lemma 2.2, $\widehat{\lambda}(t_n,x_n,u_n)(u_n-1)\to 0$. We conclude that $L(t_n,x_n,u_n)\to ct=L(t,x,1)$.

We consider the continuity at $(t,x,0)$, where $(t,x)\in(0,\infty)\times(-\infty,0]$. Suppose $(t_n,x_n)\to(t,x)$ and $u_n\downarrow 0$. If $x<0$, then $x_n<0$ for all n large enough, so we apply Lemma 2.2 part 8 to obtain that $\widehat{\lambda}(t_n,x_n,u_n)\to-\infty$ and $\widehat{\lambda}(t_n,x_n,u_n)u_n\to 0$. It follows that $H(t_n,x_n,\widehat{\lambda}(t_n,x_n,u_n))\to\log(1-e^{-ct})$ and that $L(t_n,x_n,u_n)\to-\log(1-e^{-ct})=L(t,x,0)$. Suppose that $x=0$. Since $H_\lambda(t_n,x_n,0)=e^{-ct_n}-c(x_n\vee 0\wedge t_n)\to e^{-ct}$ as $n\to\infty$, the equation $u_n=H_\lambda(t_n,x_n,\widehat{\lambda}(t_n,x_n,u_n))$ implies by part 1 of Lemma 2.2 that $\widehat{\lambda}(t_n,x_n,u_n)<0$ for all n large enough. By Lemma 2.2 part 6, $\widehat{\lambda}(t_n,x_n,u_n)u_n\to 0$ as $n\to\infty$. By the lefthand inequality of part 7 and by part 8 of Lemma 2.2, $\widehat{\lambda}(t_n,x_n,u_n)\to-\infty$. The righthand inequality of Lemma 2.2 part 7 implies that $c(x_n\vee 0\wedge t_n)e^{-\widehat{\lambda}(t_n,x_n,u_n)}\to 0$ and by (2.1), $H(t_n,x_n,\widehat{\lambda}(t_n,x_n,u_n))\to\log(1-e^{-ct})$, so $L(t_n,x_n,u_n)\to-\log(1-e^{-ct})=L(t,x,0)$.

Finally, $\widehat{\lambda}(t,x,1-)=\infty$ by part 3 of Lemma 2.2 and $\widehat{\lambda}(t,0,0+)=-\infty$ by part 8 of Lemma 2.2. □

The next lemma deals with the differentiability properties of $L(t,x,u)$.

Lemma 2.4. *The functions $\widehat{\lambda}(t,x,u)$ and $L(t,x,u)$ are continuously differentiable in (t,x,u) on the set $\{(t,x,u):0<x<t,\ u<1\}$. For these (t,x,u), the following identities hold*

$$\begin{aligned}\widehat{\lambda}(t,x,u)&=L_u(t,x,u),\\L_t(t,x,u)&=-H_t(t,x,\hat{\lambda}(t,x,u)),\\L_x(t,x,u)&=-H_x(t,x,\hat{\lambda}(t,x,u)).\end{aligned}$$

Also,

$$|L_t(t,x,u)| \le c + \frac{ce^{-ct}}{1-e^{-ct}},$$
$$L_x(t,x,u) \le c,$$
$$|L_x(t,x,u)| \le c + \frac{c}{1-e^{-ct}}\frac{e^{-ct}+|u|}{cx}.$$

The functions $L_t(t,x,u)$ and $L_x(t,x,u)$ can be continuously extended to the set $\{(t,x,u) : 0 < x \le t, u \le 1\}$. In addition, for every interval $[a,b] \subset \mathbf{R}_+$ with $a > 0$ there exist $C_1 > 0$ and $C_2 > 0$ such that $L_x(t,x,0) \le -C_1/\sqrt{x}$ for all $x \in (0, C_2]$ and $t \in [a,b]$.

Proof. Given $(t,x,u) \in (0,\infty) \times (0,\infty) \times (-\infty,1)$, we have by Lemma 2.1 that $H_{\lambda\lambda}(t,x,\widehat{\lambda}(t,x,u)) > 0$. Since $u = H_\lambda(t,x,\widehat{\lambda}(t,x,u))$ and, by Lemma 2.1, $H_\lambda(t,x,\lambda)$ is continuously differentiable in (t,x,λ) provided $0 < x < t$ and $\lambda \in \mathbf{R}$, by the implicit function theorem the function $\widehat{\lambda}(t,x,u)$ is continuously differentiable in (t,x,u) if $0 < x < t$ and $u < 1$. It follows that $L(t,x,u)$ is also continuously differentiable. The equations $L_t(t,x,u) = -H_t(t,x,\hat{\lambda}(t,x,u))$, $L_x(t,x,u) = -H_x(t,x,\hat{\lambda}(t,x,u))$ and $L_u(t,x,u) = \widehat{\lambda}(t,x,u)$ follow in a standard fashion by differentiating the equation $L(t,x,u) = \widehat{\lambda}(t,x,u)u - H(t,x,\widehat{\lambda}(t,x,u))$ and using that $u = H_\lambda(t,x,\widehat{\lambda}(t,x,u))$, cf. [**4**, page 138, Lemma 4.27].

By the definition of $H(t,x,\lambda)$ in (2.1), for $0 < x < t$ and $u < 1$,

$$\begin{aligned} L_t(t,x,u) &= -H_t(t,x,\widehat{\lambda}(t,x,u)) \\ &= \frac{c(e^{\hat{\lambda}(t,x,u)}-1)e^{-ct}}{(e^{\hat{\lambda}(t,x,u)}-1)e^{-ct} + e^{c(e^{-\hat{\lambda}(t,x,u)}-1)x}}. \end{aligned} \tag{2.7}$$

It follows that

$$|L_t(t,x,u)| \le c + \frac{ce^{-ct}}{1-e^{-ct}}.$$

Also,

$$\begin{aligned} L_x(t,x,u) &= -H_x(t,x,\widehat{\lambda}(t,x,u)) \\ &= -\frac{c(e^{-\hat{\lambda}(t,x,u)}-1)e^{c(e^{-\hat{\lambda}(t,x,u)}-1)x}}{(e^{\hat{\lambda}(t,x,u)}-1)e^{-ct} + e^{c(e^{-\hat{\lambda}(t,x,u)}-1)x}}. \end{aligned} \tag{2.8}$$

If $\lambda \geq 0$, then $(e^{\lambda}-1)e^{-ct} + e^{c(e^{-\lambda}-1)x} \geq e^{c(e^{-\lambda}-1)x}$, so that $0 \leq L_x(t,x,u) \leq c$ if $\widehat{\lambda}(t,x,u) \geq 0$. If $\lambda < 0$, then

$$\frac{c(e^{-\lambda}-1)e^{c(e^{-\lambda}-1)x}}{(e^{\lambda}-1)e^{-ct}+e^{c(e^{-\lambda}-1)x}} \leq \frac{ce^{-\lambda}}{(e^{\lambda}-1)e^{-ct}e^{-c(e^{-\lambda}-1)x}+1} \leq \frac{c\,e^{-\lambda}}{1-e^{-ct}}.$$

Hence, if $\widehat{\lambda}(t,x,u) < 0$, then

$$-\frac{c}{1-e^{-ct}}\,e^{-\hat{\lambda}(t,x,u)} \leq L_x(t,x,u) \leq 0.$$

By part 4 of Lemma 2.2, $e^{-\hat{\lambda}(t,x,u)} \leq (e^{-ct}+|u|)/(cx)$. Combining the cases $\widehat{\lambda}(t,x,u) \geq 0$ and $\widehat{\lambda}(t,x,u) < 0$, we obtain that

$$|L_x(t,x,u)| \leq c + \frac{c}{1-e^{-ct}}\frac{e^{-ct}+|u|}{cx}.$$

Let $(t_n,x_n,u_n) \to (t,x,1)$ as $n \to \infty$, where $0 < x_n < t_n$, $u_n < 1$ and $t > 0$. By Lemma 2.2 part 3, $\widehat{\lambda}(x_n,t_n,u_n) \to \infty$. By (2.7) and (2.8), $L_t(t_n,x_n,u_n) \to c$ and $L_x(t_n,x_n,u_n) \to 0$, which shows that $L_t(t,x,u)$ and $L_x(t,x,u)$ extend continuously to $\{(t,x,u) : 0 < x \leq t,\ u \leq 1\}$.

Since $H_{\lambda}(t,x,0) = e^{-ct} - cx > 0$ for all x small enough, $\widehat{\lambda}(t,x,0) < 0$ for $t \in [a,b]$ and all $x < e^{-bc}/c$. By (2.8), $L_x(t,x,0) \leq -c(e^{-\hat{\lambda}(t,x,u)} - 1)$. By Lemma 2.2 part 7, for all $t \in [a,b]$ and all $x \in (0,(e^{-bc}/c) \wedge a]$, $-\widehat{\lambda}(t,x,0) \geq -(ct+1)/2 - (\log x)/2 - (\log c)/2$. Hence, $L_x(t,x,0) \leq -\sqrt{c}e^{-(ct+1)/2}/\sqrt{x} - c$, so one can take $C_1 = \sqrt{c}e^{-(cb+1)/2}/2$ and $C_2 = (e^{-cb+1}/(4c)) \wedge (e^{-bc}/c) \wedge a$. □

Remark 2.1. One can show, in addition, that $\lim_{\substack{x\to 0\\ u\downarrow 0}} L_x(t,x,u) = \lim_{\substack{x\to 0\\ u\downarrow 0}} L_u(t,x,u) = -\infty$ and $\lim_{u\uparrow 1} L_u(t,x,u) = \infty$.

We now concern ourselves with finding a lower bound on $\mathbf{I}(x(\cdot))$. As a consequence of Lemma 2.3, if $x(\cdot)$ is absolutely continuous and $x(0) = 0$, then

$$\mathbf{I}(x(\cdot)) = \int_0^{\infty} L(t,x(t),\dot{x}(t))\,dt. \tag{2.9}$$

In particular, the righthand side equals ∞ if $x(\cdot)$ assumes negative values because in that case $\int_0^\infty \mathbf{1}_{\{\dot{x}(t)<0\}}(t)\mathbf{1}_{\{\dot{x}(t)<0\}}(t)\,dt > 0$, so $\int_0^\infty \mathbf{1}_{\{L(t,x(t),\dot{x}(t))=\infty\}}(t)\,dt > 0$ by Lemma 2.3.

Lemma 2.5. *Given $T > 0$, there exists $A_T > 0$ such that, for all absolutely continuous $x(\cdot)$,*

$$\int_0^T |\dot{x}(t)\log(|\dot{x}(t)|)|\,dt \leq A_T(1 + \mathbf{I}(x(\cdot))).$$

Proof. Since $\dot{x}(t) \leq 1$ almost everywhere if $\mathbf{I}(x(\cdot)) < \infty$, it suffices to prove that there exists $K(T) > 0$ such that

$$\int_0^T |\dot{x}(t)\log(|\dot{x}(t)|)|\mathbf{1}_{\{\dot{x}(t)\leq -K(T)\}}(t)\,dt \leq A'_T(1 + \mathbf{I}(x(\cdot))) \tag{2.10}$$

for some $A'_T > 0$. By Lemma 2.1, on taking into account that $\dot{x}(t) = 0$ almost everywhere on the set $\{t : x(t) = 0\}$, we have that $\dot{x}(t) = H_\lambda(t, x(t), \widehat{\lambda}(t, x(t), \dot{x}(t)))$ for almost all t such that $\dot{x}(t) < 0$. Since $H_\lambda(t, x, 0) \geq e^{-ct} - t$, by Lemma 2.2 part 1, if $K(T) \geq T$, then $\widehat{\lambda}(t, x(t), \dot{x}(t)) < 0$ almost everywhere on the set $\{t : \dot{x}(t) \leq -K(T)\}$. By Lemma 2.2 part 4, (2.1) and (2.2),

$$(-ct) \wedge \log(1 - e^{-ct}) < H(t, x(t), \widehat{\lambda}(t, x(t), \dot{x}(t))) \leq 1 + |\dot{x}(t)|,$$

which implies that $\int_0^T |H(t, x(t), \widehat{\lambda}(t, x(t), \dot{x}(t)))|\,dt < \infty$. The representation $L(t, x(t), \dot{x}(t)) = \widehat{\lambda}(t, x(t), \dot{x}(t))\dot{x}(t) - H(t, x(t), \widehat{\lambda}(t, x(t), \dot{x}(t)))$ and (2.9) yield

$$\int_0^T \widehat{\lambda}(t, x(t), \dot{x}(t))\dot{x}(t)\,\mathbf{1}_{\{\dot{x}(t)\leq -K(T)\}}(t)\,dt \\ \leq T + \int_0^T |\dot{x}(t)|\mathbf{1}_{\{\dot{x}(t)\leq -K(T)\}}(t)\,dt + \mathbf{I}(x(\cdot)).$$

By Lemma 2.2 part 5, if $\dot{x}(t) < 0$ and $\widehat{\lambda}(t, x(t), \dot{x}(t)) < 0$, then $-\widehat{\lambda}(t, x(t), \dot{x}(t)) \geq \log(-\dot{x}(t)) - ct$. It follows that

$$\int_0^T |\dot{x}(t)\log|\dot{x}(t)||\,\mathbf{1}_{\{\dot{x}(t)\leq -K(T)\}}(t)\,dt \\ \leq T + (1 + cT)\int_0^T |\dot{x}(t)|\,\mathbf{1}_{\{\dot{x}(t)\leq -K(T)\}}(t)\,dt + \mathbf{I}(x(\cdot)).$$

Suppose, in addition, that $\log K(T) > 2(1+cT)$. Then $|\dot{x}(t)\log|\dot{x}(t)|| \geq 2(1+cT)|\dot{x}(t)|$ if $\dot{x}(t) \leq -K(T)$. Inequality (2.10) follows. □

3. Solving the variational problem. The purpose of this section is a proof of the following result.

Theorem 3.1. *The infimum of* $\int_0^1 L(t, x(t), \dot{x}(t))\,dt$ *over all absolutely continuous functions* $x(\cdot)$ *such that* $x(0) = 0$ *and* $0 \leq x(t) \leq t$ *for* $t \in [0,1]$, *and* $x(1) = a \in [0,1]$ *is attained at a unique function* $x_{a,1}(\cdot)$ *which is defined by* $x_{a,1}(t) = (1+\alpha(e^{-c\alpha t}-1))t$, *where* α *is the nonnegative solution of the equation* $1+\alpha(e^{-c\alpha}-1) = a$. *In addition,*

$$\int_0^1 L(t, x_{a,1}(t), \dot{x}_{a,1}(t))\,dt = \frac{c}{2}\,(\alpha-1)^2 + (1-a)\log\alpha + a(\log a + \alpha c).$$

Remark 3.1. This theorem also delivers a solution to the problem of minimizing $\int_0^T L(t, x(t), \dot{x}(t))\,dt$ over nonnegative absolutely continuous $x(\cdot)$ such that $x(t) \leq t$ for $t \in [0,T]$, $x(0) = 0$ and $x(T) = a$, where $T > 0$ and $a \in [0,T]$. This can be seen by a scaling argument as follows. If we indicate the dependence of L on c explicitly by writing $L^{(c)}(t, x, u)$, then $\int_0^T L^{(c)}(t, x(t), \dot{x}(t))\,dt = T\int_0^1 L^{(cT)}(t, y(t), \dot{y}(t))\,dt$, where $y(t) = x(tT)/T$. Thus, $x_{a,T}(t) = Tx_{a/T,1}(t/T)$ is the unique minimizer and the minimum equals $T((cT/2)(\alpha_T-1)^2 + (1-a/T)\log\alpha_T + (a/T)(\log(a/T)+\alpha_T c))$, where $1+\alpha_T(e^{-cT\alpha_T}-1) = a/T$.

The proof of Theorem 3.1 proceeds through a number of lemmas. The next lemma describes some properties of the trajectory of interest.

Lemma 3.1. *Given* $0 \leq t_1 < t_2$ *and* $a \leq 1$, *there exists a unique* $\alpha \in \mathbf{R}_+$ *such that* $1+\alpha(e^{-ct_1}\,e^{-c\alpha(t_2-t_1)}-1) = a$; $\alpha = 0$ *if and only if* $a = 1$. *For* $x(t) = (1+\alpha(e^{-ct_1}e^{-c\alpha(t-t_1)}-1))(t-t_1)$ *with* $t_1 \leq t \leq t_2$ *and* $\alpha > 0$, *the supremum in the expression* (2.4) *for* $L(t, x(t), \dot{x}(t))$ *is attained at* $\widehat{\lambda}(t) = \log(1+\,e^{ct_1}\,e^{c\alpha(t-t_1)}(1-\alpha)/\alpha)$. *In addition, for* $s \in [0, t_2-t_1]$,

$$\int_{t_1}^{t_1+s} L(t, x(t), \dot{x}(t))\,dt = \frac{cs^2}{2}\,(\alpha-1)^2 + s\log\alpha + s(1+\alpha(e^{-ct_1-\alpha cs}-1)) \times \log\left(e^{ct_1+\alpha cs}\left(\frac{1}{\alpha}-1\right)+1\right).$$

Proof. With $x = c\alpha(t_2 - t_1)$, the equation for α takes the form $xe^{-x}e^{-ct_1} = x + c(t_2 - t_1)(a-1)$. Since $a \le 1$, the latter equation has a unique nonnegative solution. One checks that $\dot{x}(t) = H_\lambda(t, x(t), \widehat{\lambda}(t))$. The expression for the integral is verified by differentiation. □

Remark 3.2. An explanation of how $\widehat{\lambda}(t)$ has been found is given in the proof of Lemma 3.3.

We now address the existence of a minimizer.

Lemma 3.2. *The infimum of $\int_0^1 L(t, x(t), \dot{x}(t))\,dt$ over all absolutely continuous functions $x(\cdot)$ such that $x(0) = 0$, $0 \le x(t) \le t$ for $t \in [0,1]$, and $x(1) = a \in [0,1]$ is finite and is attained.*

Proof. The infimum is finite because $\int_0^1 L(t, x_{a,1}(t), \dot{x}_{a,1}(t))\,dt < \infty$ by Lemma 3.1. Let $\mathbf{C}[0,1]$ denote the Banach space of continuous real-valued functions $x(\cdot)$ defined on $[0,1]$ endowed with uniform norm. We define a functional J on $\mathbf{C}[0,1]$ by letting $J(x(\cdot)) = \int_0^1 L(t, x(t), \dot{x}(t))\,dt$ if $x(\cdot)$ is absolutely continuous, $x(0) = 0$, and $0 \le x(t) \le t$ for $t \in [0,1]$, and by letting $J(x(\cdot)) = \infty$ otherwise. We will prove that J is lower compact, i.e., the sets $\{x(\cdot) : J(x(\cdot)) \le q\}$ are compact for all $q \ge 0$. That will imply that $J(x(\cdot))$ attains infima on closed sets, so $\inf_{x\in\mathbf{C}[0,1]:x(0)=0,x(1)=a} J(x(\cdot))$ is attained.

Since $L(t, x, u)$ is the convex conjugate of $H(t, x, \lambda)$ in the last variable (see (2.4)), we have by Young's inequality that $\lambda u \le H(t, x, R\lambda)/R + L(t, x, u)/R$ for arbitrary $R > 0$. We let $\operatorname{sign}(x) = 1$ if $x \ge 0$ and $\operatorname{sign}(x) = -1$ if $x < 0$. Given $x(\cdot)$ such that $J(x(\cdot)) < \infty$, we have recalling (2.1) that, for $0 \le s < t \le 1$,

$$\begin{aligned} |x(t) - x(s)| &\le \int_0^1 \mathbf{1}_{[s,t]}(r)|\dot{x}(r)\,dr \\ &\le \frac{1}{R}\int_0^1 H(r, x(r), R\,\operatorname{sign}(\dot{x}(r))\mathbf{1}_{[s,t]}(r))\,dr \\ &\quad + \frac{1}{R}\int_0^1 L(r, x(r), \dot{x}(r))\,dr \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{R}\int_s^t H(r, x(r), R)\mathbf{1}_{\{\dot{x}(r)\geq 0\}}(r)\,dr \\
&\quad + \frac{1}{R}\int_s^t H(r, x(r), -R)\mathbf{1}_{\{\dot{x}(r)<0\}}(r)\,dr + \frac{J(x(\cdot))}{R} \\
&\leq (t-s) + \frac{e^R c}{R}\int_s^t x(r)\,dr + \frac{J(x(\cdot))}{R}.
\end{aligned}
$$

Since R is arbitrary and $0 \leq x(t) \leq t$, the elements of the set $\{x(\cdot) \in \mathbf{C}[0,1] : J(x(\cdot)) \leq q\}$ are uniformly equicontinuous. Since $x(0) = 0$ on that set, an application of the Arzela-Ascolli theorem shows that the set in question is relatively compact in $\mathbf{C}[0,1]$.

It remains to show that J is lower semicontinuous. Suppose $x_n(\cdot) \to x(\cdot)$ in $\mathbf{C}[0,1]$ and $J(x_n(\cdot)) < \infty$. Let Λ_0 represent the set of $\mathbf{R}$-valued simple functions on $\mathbf{R}_+$. By (2.4) and [**17**, page 460, Lemma A.2],

$$
\begin{aligned}
&\liminf_{n\to\infty}\int_0^1 L(t, x_n(t), \dot{x}_n(t))\,dt \\
&\qquad = \liminf_{n\to\infty}\sup_{\lambda(\cdot)\in\Lambda_0}\int_0^1 (\lambda(t)\dot{x}_n(t) - H(t, x_n(t), \lambda(t)))\,dt \\
&\qquad \geq \sup_{\lambda(\cdot)\in\Lambda_0}\liminf_{n\to\infty}\int_0^1 (\lambda(t)\dot{x}_n(t) - H(t, x_n(t), \lambda(t)))\,dt \\
&\qquad = \sup_{\lambda(\cdot)\in\Lambda_0}\int_0^1 (\lambda(t)\dot{x}(t) - H(t, x(t), \lambda(t)))\,dt \\
&\qquad = \int_0^1 L(t, x(t), \dot{x}(t))\,dt,
\end{aligned}
$$

where the convergence of integrals follows by $\lambda(\cdot)$ being piecewise constant, (2.1) and (2.2). □

Remark 3.3. By Lemma 2.5, the set $\{x(\cdot) : J(x(\cdot)) \leq q\}$ is relatively compact for the weak topology on $W^{1,1}(0,1)$, so one may hope to use the general theory such as [**4**, page 96, Theorem 3.23] in order to deduce the lower semicontinuity of J. However, that doesn't seem to be possible because $L(\cdot)$ is not a Carathéodory function.

Remark 3.4. The lower compactness property of $J(x(\cdot))$ is also a byproduct of Theorem 1.1.

As follows from Lemma 2.3 and Lemma 2.4, the integrand $L(t,x,u)$ does not meet the standard requirements for a minimizer to $\int_0^1 L(t,x(t),\dot{x}(t))\,dt$ over absolutely continuous $x(\cdot)$ such that $x(0)=0$ and $x(1)=a$, where $a\in[0,1]$, to satisfy the Euler-Lagrange equation, see, e.g., [**4**, page 125, Theorem 4.12]. Nevertheless, we will solve the equation and show that it delivers the minimum. The next lemma provides the essential stepping stone toward proving the optimality of the trajectory appearing in Theorem 3.1. Let us denote $\Sigma=\{(t,x): t>0, 0<x<t\}$.

Lemma 3.3. *If $\widehat{x}(\cdot)$ is a trajectory that minimizes $\int_{t_1}^{t_2} L(t,x(t),\dot{x}(t))dt$, where $0\le t_1<t_2$, among all absolutely continuous functions $x(\cdot)$ such that $(t,x(t))\in\Sigma$ for $t\in(t_1,t_2)$, $x(t_1)=0$, and $x(t_2)=a\in[0,t_2-t_1]$, then*

$$\widehat{x}(t)=(1+\widehat{\alpha}(e^{-ct_1}e^{-c\hat{\alpha}(t-t_1)}-1))(t-t_1),$$

where $\widehat{\alpha}$ is the unique nonnegative solution of the equation

$$1+\widehat{\alpha}(e^{-ct_1}\,e^{-c\hat{\alpha}(t_2-t_1)}-1)=\frac{a}{t_2-t_1}.$$

Proof. If $a=t_2-t_1$, then there is a unique trajectory such that $x(t_1)=0$, $x(t_2)=a$ and $\int_{t_1}^{t_2} L(t,x(t),\dot{x}(t))\,dt<\infty$, which is $x(t)=t-t_1$. It satisfies the requirements with $\widehat{\alpha}=0$. We now assume that $a<t_2-t_1$. Consider t_1' and t_2' with $t_1<t_1'<t_2'<t_2$. Then the piece of $\widehat{x}(t)$ for $t\in[t_1',t_2']$ is an optimal trajectory which belongs to Σ and has $(t_1',\widehat{x}(t_1'))$ as the initial point and $(t_2',\widehat{x}(t_2'))$, as the terminal point. Let $x_1'=\widehat{x}(t_1')$ and $x_2'=\widehat{x}(t_2')$. We note that by Lemma 2.3 we must have that $\dot{\widehat{x}}(t)\le 1$ almost everywhere. Letting $u(t)=\dot{x}(t)$, we reformulate the variational problem of minimizing $\int_{t_1'}^{t_2'} L(t,x(t),\dot{x}(t))\,dt$ as the Lagrange optimal control problem of minimizing $\int_{t_1'}^{t_2'} L(t,x(t),u(t))\,dt$ over absolutely continuous functions $x(\cdot)$ and measurable functions $u(\cdot)$ subject to the constraints $\dot{x}(t)=u(t)$, $x(t_1')=x_1'$, $x(t_2')=x_2'$ and $\dot{u}(t)\in(-\infty,1]$ almost everywhere. We apply Pontryagin's maximum principle, as in [**3**, pages 196, 197], with $A=\{(t,x): t_1'/2\le t\le 2t_2', \min_{s\in[t_1',t_2']}\widehat{x}(s)/2\le x\le t\}$ and $U=(-\infty,1]$. By Lemma 2.4, the function $L(t,x,u)$ is continuous on $M=A\times U$ together with L_x and L_t. By [**3**, page 197, Theorem 5.1.i], if $(x'(t), t_1'\le t\le t_2')$ is an optimal trajectory which belongs

to the interior of A and $(u'(t), t_1' \leq t \leq t_2')$ is the associated optimal control, then either there exists an absolutely continuous function $\lambda'(\cdot)$ such that the following equations hold almost everywhere

$$\dot{x}'(t) = u'(t), \tag{3.1a}$$

$$\dot{\lambda}'(t) = L_x(t, x'(t), u'(t)), \tag{3.1b}$$

and

$$\lambda'(t)u'(t) - L(t, x'(t), u'(t)) = \sup_{u \in U}(\lambda'(t)u - L(t, x'(t), u)), \tag{3.1c}$$

which occurs if λ_0 in (5.1.3) [**3**, page 197] is positive, or $u(t) = 1$ and $\dot{x}'(t) = 1$ almost everywhere if $\lambda_0 = 0$. The latter function $x'(\cdot)$ does not satisfy the requirements that $x'(t_1') = x_1'$ and $x'(t_2') = x_2'$ if t_1' and t_2' are close enough to t_1 and t_2, respectively, because $a < t_2 - t_1$, so we leave it out of consideration.

We note that Theorem 4.2.i [**3**, page 162], which is invoked for the proof of Theorem 5.1.i there, requires that the optimal control be bounded which we do not know a priori. However, Remark 5 [**3**, page 167] allows us to incorporate unbounded controls by checking condition (S). It stipulates that there exist a nonnegative Lebesgue integrable on $[t_1', t_2']$ function $S(t)$ and $\gamma > 0$ such that for all $t \in [t_1', t_2']$ and all $(\widetilde{t}, \widetilde{x}) \in A$ subject to the requirements that $|\widetilde{t} - t| \leq \gamma$ and $|\widetilde{x} - x'(t)| \leq \gamma$ the following holds

$$|L_t(\widetilde{t}, \widetilde{x}, u'(t))| + |L_x(\widetilde{t}, \widetilde{x}, u'(t))| \leq S(t).$$

Since by Lemma 2.4,

$$|L_t(t, x, u)| \leq c + \frac{ce^{-ct}}{1 - e^{-ct}}$$

and

$$|L_x(t, x, u)| \leq c + \frac{c}{1 - e^{-ct}} \frac{e^{-ct} + |u|}{cx},$$

we can take $\gamma = \min_{t\in[t_1',t_2']}(x'(t)\wedge(t-x'(t)))/2$ and

$$S(t) = c + \frac{ce^{-ct}}{1-e^{-ct}} + c + \frac{c}{1-e^{-ct}}\,\frac{e^{-ct}+|u'(t)|}{c\gamma}.$$

We have thus checked the applicability of Pontryagin's maximum principle in our setting.

Since $L(t,x,u)$ is the convex conjugate of $H(t,x,\lambda)$ in λ (see (2.4)), we have that the righthand side of (3.1c) equals $H(t,x'(t),\lambda'(t))$, so $\sup_\lambda(\lambda u'(t) - H(t,x'(t),\lambda'(t))$ is attained at $\lambda = \lambda'(t)$, which implies that $\lambda'(t) = \widehat{\lambda}(t,x'(t),u'(t))$ almost everywhere and $u'(t) = H_\lambda(t,x'(t),\lambda't))$ almost everywhere. It follows that $u(t) < 1$ almost everywhere and that $L_x(t,x'(t),u'(t)) = -H_x(t,x'(t),\lambda'(t))$ (for instance, by Lemma 2.4). Equations (3.1a) and (3.1b) take the canonical form

$$\dot{x}'(t) = H_\lambda(t,x'(t),\lambda'(t)),$$
$$\dot{\lambda}'(t) = -H_x(t,x'(t),\lambda'(t)).$$

Consequently,

(3.2a)
$$\dot{x}'(t) = \frac{e^{\lambda'(t)}e^{-ct} - ce^{c(e^{-\lambda'(t)}-1)x'(t)}e^{-\lambda'(t)}x'(t)}{(e^{\lambda'(t)}-1)e^{-ct} + e^{c(e^{-\lambda'(t)}-1)x'(t)}},$$

(3.2b)
$$\dot{\lambda}'(t) = -\frac{ce^{c(e^{-\lambda'(t)}-1)x'(t)}(e^{-\lambda'(t)}-1)}{(e^{\lambda'(t)}-1)e^{-ct} + e^{c(e^{-\lambda'(t)}-1)x'(t)}}.$$

Let us introduce

(3.3a) $$y(t) = e^{c(e^{-\lambda'(t)}-1)x'(t)}$$

and

(3.3b) $$\mu(t) = (e^{\lambda'(t)}-1)e^{-ct}.$$

We note that $y(t) > 0$ and $y(t)+\mu(t) > 0$ (see (2.2) for the latter). Equations (3.2a) and (3.2b) imply that

(3.4a) $$\dot{y}(t) = -\frac{cy(t)\mu(t)}{y(t)+\mu(t)},$$

(3.4b) $$\dot{\mu}(t) = -\frac{c\mu(t)^2}{y(t)+\mu(t)}.$$

The righthand side of (3.4b) is Lipshitz-continuous in $\mu(t)$ in a small enough neighborhood of t_1', which implies that if $\mu(t_1') = 0$, then that equation has the unique solution $\mu(t) = 0$ in such a neighborhood. By (3.4a), $y(\cdot)$ does not vary over that neighborhood, so we can apply this argument repeatedly to conclude that $\mu(t) = 0$ for all t if $\mu(t_1') = 0$. By (3.3b), $\lambda'(t) = 0$ and, by (3.2a), $\dot{x}'(t) = e^{-ct} - cx'(t)$, so

$$x'(t) = (x_1' + t - t_1')e^{-c(t-t_1')}. \tag{3.5}$$

It is an admissible trajectory provided $(x_1' + t_2' - t_1')e^{-c(t_2'-t_1')} = x_2'$. Since $x_1' \to 0$ and $x_2' \to a$ as $t_1' \to t_1$ and $t_2' \to t_2$, this condition implies that $(t_2 - t_1)e^{-c(t_2-t_1)} = a$. We then obtain the assertion of the lemma with $\widehat{\alpha} = 1$.

We now solve the equations assuming $\mu(t_1') \neq 0$. By (3.4b), $\mu(t)$ is monotonically nonincreasing, so $\mu(t)$ maintains its sign provided $\mu(t_1') < 0$. Suppose $\mu(t_1') > 0$, and let $\breve{t} = \inf\{t \geq t_1 : \mu(t) \leq 0\} \leq \infty$. Since $y(t) > 0$, we have that $\dot{\mu}(t) \geq -c\mu(t)$ for $t < \breve{t}$, so that $\mu(t) \geq \mu(t_1')\exp(-c(t - t_1'))$, which implies that $\breve{t} = \infty$. Thus, $\mu(t)$ also maintains its sign if $\mu(t_1') > 0$. Since $\mu(t) \neq 0$ and $y(t) \neq 0$ for all t, we can divide the righthand sides of (3.4a) and (3.4b) by $y(t)$ and $\mu(t)$, respectively, to deduce that $\dot{y}(t)/y(t) = \dot{\mu}(t)/\mu(t)$. Therefore, $y(t) = K\mu(t)$ where $K \neq 0$. Since $y(t) + \mu(t) > 0$ and $y(t) > 0$, we also have that $(K+1)/K > 0$, so either $K < -1$ or $K > 0$. By (3.4b), $\mu(t) = \mu(t_1')\exp(-(c/(1+K))(t - t_1'))$. With $\alpha' = K/(K+1)$ and $\beta' = y(t_1')$, we deduce with the aid of (3.3a) and (3.3b) that

(3.6a)
$$x'(t) = \left(1 + \alpha'\left(\frac{e^{-ct}}{\beta'}e^{-c(\alpha'-1)(t-t_1')} - 1\right)\right)\left(t - t_1' - \frac{\log\beta'}{c(1-\alpha')}\right),$$

(3.6b)
$$\lambda'(t) = \log\left(1 + \beta'\frac{1-\alpha'}{\alpha'}e^{ct}e^{c(\alpha'-1)(t-t_1')}\right),$$

where $\alpha' > 0$, $\alpha' \neq 1$, and $\beta' > e^{-cx_1'}$. Substituting $t = t_1'$ and $t = t_2'$

in (3.6a), we obtain that

(3.7a)
$$x_1' = \left(1 + \alpha'\left(\frac{e^{-ct_1'}}{\beta'} - 1\right)\right)\frac{\log\beta'}{c(\alpha'-1)},$$
(3.7b)
$$x_2' = \left(1 + \alpha'\left(\frac{e^{-ct_1'}}{\beta'}\, e^{-c\alpha'(t_2'-t_1')} - 1\right)\right)\left(t_2' - t_1' - \frac{\log\beta'}{c(1-\alpha')}\right).$$

Besides, noting that $y(t_1') = (\alpha'/(1-\alpha'))\mu(t_1')$, we can see that if $\alpha' > 1$, then $\mu(t_1') < 0$, so by (3.3a) and (3.3b) $\beta' = y'(t_1') > 1$. Analogously, if $\alpha' < 1$, then $\beta' < 1$. Therefore, $\log\beta'/(1-\alpha') < 0$.

We now let $t_1' \to t_1$ and $t_2' \to t_2$. Since $x_1' = \widehat{x}(t_1')$ and $x_2' = \widehat{x}(t_2')$, we have that $x_1' \to 0$ and $x_2' \to a$. We show that $\beta' \to 1$. The bound $\beta' > e^{-cx_1'}$ implies that $\liminf\beta' \geq 1$. Let β'' be a subsequence with $\lim\beta'' = \limsup\beta'$. The inequality $\log\beta'/(1-\alpha') < 0$ implies that $t_2' - t_1' - (\log\beta')/(c(1-\alpha')) > t_2' - t_1'$. Therefore, the lefthand multiplier on the right of (3.7b) is nonnegative, which yields

$$1 + \alpha'\left(\frac{e^{-ct_1'}}{\beta'} - 1\right) \geq (\alpha'-1)(e^{c\alpha'(t_2-t_1)} - 1). \tag{3.8}$$

Let α'' be the subsequence associated with β''. If $\alpha'' > 1$ eventually, then by (3.8), the $(1 + \alpha''(e^{-ct_1'}/\beta'' - 1))/(\alpha''-1)$ are eventually bounded from below by a positive quantity, so the convergence of the righthand side of (3.7a) to zero yields the convergence $\lim\log\beta'' = 0$. If $\alpha'' < 1$ infinitely often, then $\beta'' < 1$ infinitely often (recall that $\log\beta'/(1-\alpha') < 0$). It follows that $\lim\beta'' \leq 1$. The convergence $\beta' \to 1$ has been proved.

By (3.8), the net α' is bounded. Let $\widehat{\alpha}$ represent a limit point. The convergence of the righthand side of (3.7a) to zero implies that $\log\beta'/(1-\alpha') \to 0$. By (3.7b),

$$1 + \widehat{\alpha}(e^{-ct_1}e^{-c\hat{\alpha}(t_2-t_1)} - 1) = \frac{a}{t_2 - t_1},$$

which has a unique positive solution. It follows that $\alpha' \to \widehat{\alpha}$. Since $x'(\cdot) = \widehat{x}(\cdot)$ on $[t_1', t_2']$, by (3.6a), $\widehat{x}(t) = (1 + \widehat{\alpha}(e^{-ct_1}e^{-c\hat{\alpha}(t-t_1)} - 1))(t - t_1)$. □

In the next two lemmas we eliminate from the list of potential minimizers trajectories that hit zero before reaching the final destination. We first show that a trajectory that spends nonzero time at zero cannot be optimal.

Lemma 3.4. *If $x(\cdot)$ is a trajectory with $x(0) = 0$ and $0 \le x(t) \le t$ for $t \in [0,1]$ such that $\int_0^1 L(t, x(t), \dot{x}(t))\,dt < \infty$ and $\int_0^1 \mathbf{1}_{\{x(t)=0\}}(t)\,dt > 0$, then there exists a trajectory $\widetilde{x}(\cdot)$ with the same initial and terminal points such that $\int_0^1 L(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\,dt < \int_0^1 L(t, x(t), \dot{x}(t))\,dt$.*

Proof. The trajectory that "does better than $x(\cdot)$" will be obtained by applying a small perturbation to $x(\cdot)$. We denote $\gamma = \int_0^1 \mathbf{1}_{\{x(t)=0\}}(t)\,dt > 0$. Let $t_1 = \inf\{t > 0 : x(t) < t\}$. Let t_2 represent a Lebesgue point of $x(\cdot)$ in the interval $((1-\gamma/2) \vee t_1, 1]$. Given $\epsilon > 0$, let $\tau_\epsilon = \inf\{t > t_1 : x(t) < t - \epsilon\}$. Obviously, $\tau_\epsilon > \epsilon$. We introduce a variation $x_\epsilon(\cdot)$ of $x(\cdot)$ by letting $x_\epsilon(0) = 0$ and

$$\begin{aligned}\dot{x}_\epsilon(t) = {}& \mathbf{1}_{[0,(t_1+\epsilon)\wedge\tau_\epsilon)}(t) + \dot{x}(t)\mathbf{1}_{[(t_1+\epsilon)\wedge\tau_\epsilon, t_2-\epsilon)}(t) \\ & + \frac{1}{\epsilon}\left(x(t_2) - x_\epsilon(t_2-\epsilon)\right)\mathbf{1}_{[t_2-\epsilon,t_2)}(t) + \dot{x}(t)\mathbf{1}_{[t_2,1]}(t).\end{aligned}$$

For ϵ small enough, $0 < x_\epsilon(t) \le t$ for $t \in (0, t_2)$, $x_\epsilon(t) = x(t) = t$ for $t \in [0, t_1]$, $x(t) < x_\epsilon(t) \le x(t) + \epsilon$ for $t \in (t_1, t_2 - \epsilon]$, and $x_\epsilon(t) = x(t)$ for $t \in [t_2, 1]$. Since $L_x(t, x, u) \le c$ for $x \in (0, t)$ by Lemma 2.4 and $\dot{x}_\epsilon(t) = \dot{x}(t)$ for $t \in [(t_1+\epsilon)\wedge\tau_\epsilon, t_2-\epsilon)$, we have that for almost all $t \in [(t_1+\epsilon)\wedge\tau_\epsilon, t_2-\epsilon)$,

$$L(t, x_\epsilon(t), \dot{x}_\epsilon(t)) - L(t, x(t), \dot{x}(t)) \le c\epsilon. \tag{3.9}$$

Since, by Lemma 2.4, $L_x(t, x, 0) \le -C_1/\sqrt{x}$ for all $t \in [\epsilon, T]$ and all $x \in (0, C_2]$, where $C_1 > 0$ and $C_2 > 0$, we have, recalling that $\dot{x}(t) = 0$ almost everywhere on the set $\{t : x(t) = 0\}$, that for all ϵ small enough, almost everywhere on the set $\{t : x(t) = 0\} \cap [(t_1+\epsilon)\wedge\tau_\epsilon, t_2-\epsilon)$,

$$L(t, x_\epsilon(t), \dot{x}_\epsilon(t)) - L(t, x(t), \dot{x}(t)) \le -2C_1\sqrt{\epsilon}. \tag{3.10}$$

By Lemma 2.3, for $t \in [t_1, (t_1+\epsilon)\wedge\tau_\epsilon)$,

$$L(t, x_\epsilon(t), \dot{x}_\epsilon(t)) = ct. \tag{3.11}$$

On the interval $[t_2-\epsilon,t_2)$, the derivative of $x_\epsilon(\cdot)$ equals $(x(t_2)-x_\epsilon(t_2-\epsilon))/\epsilon$. Since the $(x(t_2)-x(t_2-\epsilon))/\epsilon$ converge to a finite limit as $\epsilon\to 0$ and $x(t_2-\epsilon)<x_\epsilon(t_2-\epsilon)\le x(t_2-\epsilon)+\epsilon$, the net $(x(t_2)-x_\epsilon(t_2-\epsilon))/\epsilon$ is bounded for all $\epsilon>0$ small enough. If $x(t_2)>0$, then the $\inf_{t\in[t_2-\epsilon,t_2]}x_\epsilon(t)$ are bounded from below for all small enough $\epsilon>0$ by a positive number, so the function $L(t,x_\epsilon(t),\dot{x}_\epsilon(t))$ is essentially bounded on $[t_2-\epsilon,t_2]$ uniformly in ϵ because $L(t,x,u)$ is continuous on $(0,\infty)\times(0,\infty)\times(-\infty,1]$ by Lemma 2.3. In that case, we have that, for some $B>0$, for almost all $t\in[t_2-\epsilon,t_2]$

$$L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\le B. \tag{3.12}$$

Suppose $x(t_2)=0$. Let $u_\epsilon=-x_\epsilon(t_2-\epsilon)/\epsilon$. Since $x_\epsilon(t)>0$ for $t\in(0,t_2)$, by Lemma 2.3, $\widehat{\lambda}(t,x_\epsilon(t),u_\epsilon)$ is well defined on $[t_2-\epsilon,t_2]$. Since the u_ϵ are bounded for all $\epsilon>0$ small enough and $x_\epsilon(t)=(t-t_2)u_\epsilon$ for $t\in[t_2-\epsilon,t_2]$, by Lemma 2.2 part 4, there exists $C_3>0$ such that $-\widehat{\lambda}(t,x_\epsilon(t),u_\epsilon)\le C_3(1-\log((t-t_2)u_\epsilon))$ for all $t\in[t_2-\epsilon,t_2)$ and all $\epsilon>0$ small enough. Hence, $\widehat{\lambda}(t,x_\epsilon(t),u_\epsilon)u_\epsilon\le C_3(1-\log((t-t_2)u_\epsilon))|u_\epsilon|$. By Lemma 2.2 parts 2 and 4 and by (2.1), the $H(t,x_\epsilon(t),\hat{\lambda}(t,x_\epsilon(t),u_\epsilon))$ are bounded on $[t_2-\epsilon,t_2]$ uniformly in ϵ. Since

$$L(t,x_\epsilon(t),u_\epsilon)=\widehat{\lambda}(t,x_\epsilon(t),u_\epsilon)u_\epsilon-H(t,x_\epsilon(t),\widehat{\lambda}(t,x_\epsilon(t),u_\epsilon)),$$

for some $C_4>0$ and all $\epsilon>0$ small enough,

$$\int_{t_2-\epsilon}^{t_2}L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\,dt\le -C_4\epsilon\log\epsilon. \tag{3.13}$$

By (3.9)–(3.13), for all $\epsilon>0$ small enough and some $C_5>0$,

$$\int_0^1 L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\,dt\le 2c\epsilon-\left(\frac{\gamma}{2}-2\epsilon\right)2C_1\sqrt{\epsilon}-C_5\epsilon\log\epsilon + \int_0^1 L(t,x(t),\dot{x}(t))\,dt.$$

Hence, if $\epsilon>0$ is small enough, then

$$\int_0^1 L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\,dt<\int_0^1 L(t,x(t),\dot{x}(t))\,dt. \qquad \square$$

The next lemma deals with trajectories that hit zero finitely many times.

Lemma 3.5. *If $x(\cdot)$ is a trajectory such that $x(t_i) = 0$ for $i = 0, 1, \dots, k-1$ and $x(t) > 0$ for all other $t \in (0, t_k)$, where $0 = t_0 < t_1 < t_2 < \cdots < t_k \le 1$, then there exists a trajectory $\widetilde{x}(\cdot)$ with $\widetilde{x}(0) = 0$, $\widetilde{x}(t) > 0$ for $t \in (0, t_k)$, and $\widetilde{x}(t_k) = x(t_k)$ such that $\int_0^{t_k} L(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\, dt < \int_0^{t_k} L(t, x(t), \dot{x}(t))\, dt$.*

Proof. It suffices to consider the case $k = 2$. Suppose $x(\cdot)$ is optimal. Then it yields optimal ways to get from $(0,0)$ to $(t_1, 0)$ and from $(t_1, 0)$ to $(t_2, x(t_2))$ staying in Σ, so by Lemma 3.3,

$$x(t) = (1 + \alpha_1(e^{-c\alpha_1 t} - 1))t \quad \text{for } 0 \le t \le t_1$$

and

$$x(t) = (1 + \alpha_2(e^{-ct_1} e^{-c\alpha_2(t-t_1)} - 1))(t - t_1) \quad \text{for } t_1 \le t \le t_2,$$

where α_1 and α_2 are the unique positive solutions of the respective equations

$$1 + \alpha_1(e^{-c\alpha_1 t_1} - 1) = 0$$

and

$$1 + \alpha_2(e^{-ct_1} e^{-c\alpha_2(t_2-t_1)} - 1) = \frac{x(t_2)}{t_2 - t_1}.$$

We note that $x(\cdot)$ is twice continuously differentiable in a left and in a right neighbourhood of t_1 and the derivatives have finite limits at t_1, in particular, $\dot{x}(t_1-) < 0 < \dot{x}(t_1+)$. By Lemma 3.1, for $t \in [0, t_1]$,

$$\begin{aligned} L(t, x(t), \dot{x}(t)) = \log\left(1 + \frac{1-\alpha_1}{\alpha_1} e^{c\alpha_1 t}\right) & \\ \times (-c\alpha_1^2 e^{-c\alpha_1 t} t + 1 + \alpha_1(e^{-c\alpha_1 t} - 1)) & \\ + \log \alpha_1 - c(\alpha_1 - 1)t. & \end{aligned}$$

In particular, $\lim_{t\uparrow t_1} L(t,x(t),\dot{x}(t)) = \infty$.

Given $\epsilon > 0$ small enough, let $t_{1,\epsilon} = \sup\{t \le t_1 : x(t) = \epsilon\}$ and $t_{2,\epsilon} = \inf\{t \ge t_1 : x(t) = \epsilon\}$. We define $x_\epsilon(t)$ by letting $x_\epsilon(t) = x(t)$ for $t \in [0, t_{1,\epsilon}]$ and $t \in [t_{2,\epsilon}, t_2]$ and $x_\epsilon(t) = \epsilon$ for $t \in [t_{1,\epsilon}, t_{2,\epsilon}]$. We have

$$\begin{aligned}\int_0^{t_2} L(t,x(t),\dot{x}(t))\,dt \quad &- \int_0^{t_2} L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\,dt \\ &= \int_{t_{1,\epsilon}}^{t_1} L(t,x(t),\dot{x}(t))\,dt \\ &\quad + \int_{t_1}^{t_{2,\epsilon}} L(t,x(t),\dot{x}(t))\,dt \\ &\quad - \int_{t_{1,\epsilon}}^{t_{2,\epsilon}} L(t,\epsilon,0)\,dt.\end{aligned}$$

Since $L(t,x(t),\dot{x}(t)) \to \infty$ as $t \uparrow t_1$ and $L(t,\epsilon,0) \to L(t_1,0,0) = -\log(1-e^{-ct_1})$ as $t \to t_1$ and $\epsilon \to 0$ (Lemma 2.3), we conclude that, given arbitrary $B > 0$, for all $\epsilon > 0$ small enough,

$$\begin{aligned}\int_0^{t_2} L(t,x(t),\dot{x}(t))\,dt \quad &- \int_0^{t_2} L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\,dt \\ &\ge B(t_1 - t_{1,\epsilon}) \\ &\quad + (\log(1-e^{-ct_1}) - 1)(t_{2,\epsilon} - t_{1,\epsilon}).\end{aligned}$$

We have that $\epsilon/(t_1 - t_{1,\epsilon}) \to -\dot{x}(t_1-) > 0$ and that $\epsilon/(t_{2,\epsilon} - t_1) \to \dot{x}(t_1+) > 0$ as $\epsilon \to 0$. Therefore, for all $\epsilon > 0$ small enough, $t_1 - t_{1,\epsilon} \ge -\epsilon/(2\dot{x}(t_1-))$ and $t_{2,\epsilon} - t_1 \le 2\epsilon/\dot{x}(t_1+)$. We obtain that, if ϵ is small enough and B is large enough, then

$$\begin{aligned}\int_0^{t_2} L(t,x(t),\dot{x}(t))\,dt \quad &- \int_0^{t_2} L(t,x_\epsilon(t),\dot{x}_\epsilon(t))\,dt \\ &\ge (B + \log(1-e^{-ct_1}) - 1)\frac{\epsilon}{-2\dot{x}(t_1-)} \\ &\quad + (\log(1-e^{-ct_1}) - 1)\frac{2\epsilon}{\dot{x}(t_1+)}.\end{aligned}$$

If B is large enough, then the righthand side is positive for all $\epsilon > 0$. It follows that $x_\epsilon(\cdot)$ meets the requirements for $\widetilde{x}(\cdot)$ in the statement of the lemma. □

Proof of Theorem 3.1. If $a = 1$, then there is the unique trajectory $x(t) = t$ with $\int_0^1 L(t, x(t), \dot{x}(t))\, dt < \infty$ which corresponds to $\alpha = 0$. Suppose $a < 1$. According to Lemmas 3.3, 3.4 and 3.5, the proof will be complete if we show that a trajectory that assumes the value of zero infinitely many times and spends no time at zero cannot be optimal. Suppose that $x(t) = 0$ at infinitely many points t in $(0, 1)$. Pick t_1 and t_2 with $0 < t_1 < t_2 < 1$ and $x(t_1) = x(t_2) = 0$ such that $x(\cdot)$ attains zero at some $t_3 \in (t_1, t_2)$. Then the set $\{t \in (t_1, t_3) : x(t) > 0\}$ is expressed as a disjoint union of possibly empty intervals $\cup_{i=1}^{\infty}(s_i, r_i)$. Since $\sum_{i=1}^{\infty}(r_i - s_i) = t_3 - t_1$, given $\epsilon > 0$, there exists N such that $\sum_{i=N+1}^{\infty}(r_i - s_i) < \epsilon$. We note that the set $V_\epsilon = (t_1, t_3) \setminus \cup_{i=1}^{N}[s_i, r_i]$ can be expressed as a finite union of disjoint open intervals. We define $x_\epsilon(\cdot)$ as coinciding with $x(\cdot)$ on the set $U_\epsilon = \cup_{i=1}^{N}(s_i, r_i)$ and by being of the form given in Lemma 3.1 (or Lemma 3.3) for $a = 0$ on the finite collection of intervals that make up the set V_ϵ. It follows by Lemma 3.1 that, for an interval $(u, v) \in V_\epsilon$,

$$\int_u^v L(t, x_\epsilon(t), \dot{x}_\epsilon(t))\, dt = \frac{1}{2} c(\widetilde{\alpha} - 1)^2 (v - u)^2 + (v - u) \log \widetilde{\alpha},$$

where $1 + \widetilde{\alpha}(e^{-cu}\, e^{-c\tilde{\alpha}(v-u)} - 1) = 0$. The latter equation implies that $1 < \widetilde{\alpha} < 1/(1 - e^{-cu})$. Since $u \ge t_1$, we obtain that

$$\int_u^v L(t, x_\epsilon(t), \dot{x}_\epsilon(t))\, dt \le \frac{1}{2} \frac{c e^{-2ct_1}}{(1 - e^{-ct_1})^2} (v - u)^2 - (v - u) \log(1 - e^{-ct_1}).$$

Since the total length of the intervals in V_ϵ is less than ϵ, for some constant K which is a function of t_1 and c only, $\int_{V_\epsilon} L(t, x_\epsilon(t), \dot{x}_\epsilon(t))\, dt \le K\epsilon$. Therefore,

$$\int_{t_1}^{t_3} L(t, x_\epsilon(t), \dot{x}_\epsilon(t))\, dt \le \int_{t_1}^{t_3} L(t, x(t), \dot{x}(t))\, dt + K\epsilon. \tag{3.14}$$

Let $\widetilde{x}(\cdot)$ represent the function defined on $[t_1, t_3]$ as in Lemma 3.1 for $a = 0$. By Lemmas 3.3 and 3.5,

$$\int_{t_1}^{t_3} L(t, \widetilde{x}(t), \dot{\widetilde{x}}(t))\, dt \le \int_{t_1}^{t_3} L(t, x_\epsilon(t), \dot{x}_\epsilon(t))\, dt.$$

Since $\epsilon > 0$ is arbitrary, by (3.14),

$$\int_{t_1}^{t_3} L(t, \tilde{x}(t), \dot{\tilde{x}}(t))\, dt \le \int_{t_1}^{t_3} L(t, x(t), \dot{x}(t))\, dt. \tag{3.15}$$

Similarly, if $\breve{x}(\cdot)$ represents the function defined as in Lemma 3.1 on $[t_3, t_2]$ for $a = 0$, then

$$\int_{t_3}^{t_2} L(t, \breve{x}(t), \dot{\breve{x}}(t))\, dt \le \int_{t_3}^{t_2} L(t, x(t), \dot{x}(t))\, dt. \tag{3.16}$$

By Lemmas 3.3 and 3.5, for the function $\overline{x}(\cdot)$ defined on $[t_1, t_2]$ as in Lemma 3.1 for $a = 0$, we have that

$$\begin{aligned} \int_{t_1}^{t_2} L(t, \overline{x}(t), \dot{\overline{x}}(t))\, dt < \int_{t_1}^{t_3} L(t, \tilde{x}(t), \dot{\tilde{x}}(t))\, dt \\ + \int_{t_3}^{t_2} L(t, \breve{x}(t), \dot{\breve{x}}(t))\, dt. \end{aligned} \tag{3.17}$$

Putting together (3.15), (3.16) and (3.17), we conclude that

$$\int_{t_1}^{t_2} L(t, \overline{x}(t), \dot{\overline{x}}(t))\, dt < \int_{t_1}^{t_2} L(t, x(t), \dot{x}(t))\, dt.$$

Thus, $x(\cdot)$ is not optimal. □

4. Proofs of Theorem 1.1 and Corollary 1.1.

Proof of Theorem 1.1. We apply Theorem 5.4.2 [**17**, page 417] to the semimartingales $X^{(n)}(\cdot)$. Let $\mu^{(n)}(\cdot)$ represent the measure of jumps of $X^{(n)}(\cdot)$ defined by

$$\mu^{(n)}([0, t], \Gamma) = \sum_{i=1}^{\lfloor nt \rfloor} \mathbf{1}_{\{\Delta X^{(n)}(i/n) \in \Gamma \setminus \{0\}\}},$$

where $\Delta X^{(n)}(s) = X^{(n)}(s) - X^{(n)}(s-)$ and $\Gamma \in \mathcal{B}(\mathbf{R})$. Let the complete σ-algebra $\mathcal{F}_t^{(n)}$ on Ω be generated by the random variables $\alpha_{ik}^{(n)}$, where

$k = 1, 2, \dots, \lfloor nt \rfloor$ and $i < k$. The predictable measure of jumps of $X^{(n)}(\cdot)$ relative to the filtration $\mathbf{F}^{(n)} = (\mathcal{F}_t^{(n)}, t \in \mathbf{R}_+)$ is defined as

$$\nu^{(n)}([0,t],\Gamma) = \sum_{i=1}^{\lfloor nt \rfloor} \mathbf{P}(\Delta X^{(n)}(i/n) \in \Gamma \setminus \{0\} | \mathcal{F}^{(n)}_{(i-1)/n}).$$

By (1.2), the random variables $\xi^{(n)}_{i,k-1}$, where $i = 1, 2, \dots, k-1$, are $\mathcal{F}^{(n)}_{(k-1)/n}$-measurable. Since $\sum_{i=1}^{k-1} \xi^{(n)}_{i,k-1} = V^{(n)}_{k-1}$ (which follows from (1.1) and (1.2) by induction), the rightmost term of (1.1) has a binomial distribution with parameters $V^{(n)}_{k-1}$ and c_n/n when conditioned on $\mathcal{F}^{(n)}_{(k-1)/n}$. Therefore, by (1.1), for m being a nonzero integer,

$$\begin{aligned}\mathbf{P}(V^{(n)}_k - V^{(n)}_{k-1} = m | \mathcal{F}^{(n)}_{(k-1)/n}) &= \left(1 - \frac{c_n}{n}\right)^{k-1} \mathbf{1}_{\{m=1\}} \\ &\quad + \sum_{l=1}^{V^{(n)}_{k-1}} \binom{V^{(n)}_{k-1}}{l} \left(\frac{c_n}{n}\right)^l \left(1 - \frac{c_n}{n}\right)^{V^{(n)}_{k-1} - l} \mathbf{1}_{\{m=-l\}},\end{aligned}$$

so using that $X^{(n)}(t) = V^{(n)}_{\lfloor nt \rfloor}/n$, we obtain that

$$\nu^{(n)}([0,t],\Gamma) = \sum_{k=1}^{\lfloor nt \rfloor} \widetilde{\nu}^{(n)}_k \left(X^{(n)}\left(\frac{k-1}{n}\right), \Gamma\right), \tag{4.1}$$

where $\widetilde{\nu}^{(n)}_k(x,\cdot)$, for $x \in \mathbf{R}$, is a measure on $\mathcal{B}(\mathbf{R})$ given by

(4.2)
$$\begin{aligned}\widetilde{\nu}^{(n)}_k(x,\cdot) &= \left(1 - \frac{c_n}{n}\right)^{k-1} \delta_{1/n} \\ &\quad + \sum_{j=1}^{\lfloor n(x \vee 0 \wedge ((k-1)/n)) \rfloor} \binom{\lfloor n(x \vee 0 \wedge ((k-1)/n)) \rfloor}{j} \left(\frac{c_n}{n}\right)^j \\ &\quad \times \left(1 - \frac{c_n}{n}\right)^{\lfloor n(x \vee 0 \wedge ((k-1)/n)) \rfloor - j} \delta_{-j/n},\end{aligned}$$

δ_y denoting the Dirac measure at y.

Obviously, $\int_{\mathbf{R}} e^{\lambda y}\widetilde{\nu}_k^{(n)}(x,dy)$ is finite for all $\lambda \in \mathbf{R}$ and $x \in \mathbf{R}$, so condition (5.4.4) [**17**, page 415] and condition (5.4.6) [**17**, page 416] hold with $\alpha_\phi = \beta_\phi = n$, and $\widehat{\nu}^{\phi}_{k/\alpha_\phi}(\beta_\phi\, dx; u) = \widetilde{\nu}_k^{(n)}(u, dx)$. We introduce, for $t \in \mathbf{R}_+$, $x \in \mathbf{R}$ and $\lambda \in \mathbf{R}$, the analogue of $\overline{g}_s^{\phi}(\lambda; u)$ in [**17**, page 417] with $r_\phi = n$ by

$$\begin{aligned} H^{(n)}(t,x,\lambda) &= \log\left(1 + \int_{\mathbf{R}} (e^{\lambda n y} - 1)\widetilde{\nu}^{(n)}_{\lfloor nt\rfloor+1}(x,dy)\right) \\ &= \log\left((e^{\lambda}-1)\left(1-\frac{c_n}{n}\right)^{\lfloor nt\rfloor} \right. \\ &\qquad + \sum_{j=0}^{\lfloor n(x\vee 0\wedge t)\rfloor} e^{-\lambda j}\binom{\lfloor n(x\vee 0\wedge t)\rfloor}{j}\left(\frac{c_n}{n}\right)^j \\ &\qquad \left. \times \left(1-\frac{c_n}{n}\right)^{\lfloor n(x\vee 0 \wedge t)\rfloor - j}\right). \end{aligned}$$

Since by Le Cam's inequality, see, e.g., Steele [**19**], for $x \in \mathbf{R}_+$,

$$\begin{aligned} \sum_{j=0}^{\lfloor nx\rfloor}\left|\binom{\lfloor nx\rfloor}{j}\left(\frac{c_n}{n}\right)^j\left(1-\frac{c_n}{n}\right)^{\lfloor nx-j\rfloor} - \frac{(c_n\lfloor nx\rfloor/n)^j}{j!}e^{-c_n\lfloor nx\rfloor/n}\right| \\ + \sum_{j=\lfloor nx\rfloor+1}^{\infty}\frac{(c_n\lfloor nx\rfloor/n)^j}{j!}e^{-c_n\lfloor nx\rfloor/n} \le \frac{2\lfloor nx\rfloor c_n^2}{n^2}, \end{aligned}$$

recalling (2.1),

$$\lim_{n\to\infty}\sup_{s\le t}\sup_{x\in\mathbf{R}}|H^{(n)}(s,x,\lambda) - H(s,x,\lambda)| = 0,$$

which checks the condition of the convergence of integrals in the statement of Theorem 5.4.2 [**17**, page 417] with $\overline{g}_s(\lambda;u) = H(s,u,\lambda)$. We note that $\overline{g}_s(\lambda;u)$ is continuous in u by (2.1). The linear growth condition $\overline{g}_s(\lambda;u) \le \widetilde{g}_s(|\lambda|(1+|u|))$ in [**17**, page 417] is met for $\widetilde{g}_s(v) = ce^{|v|} - \log(1-e^{-cs})$.

By Theorem 5.4.2 [**17**, page 417], the $X^{(n)}(\cdot)$ obey the LDP with large deviation function $\mathbf{I}$ for the Skorohod J_1-topology on the space of rightcontinuous functions on $\mathbf{R}_+$ with lefthand limits provided any solution $\mathbf{\Pi}$ of the maxingale problem $(0, G)$ as defined [**17**, page 202] is of the form $\mathbf{\Pi}(x(\cdot)) = \exp(-\mathbf{I}(x(\cdot))$, where the cumulant $G = (G_t(\lambda; x(\cdot)))$ is given by

$$G_t(\lambda; x(\cdot)) = \int_0^t H(s, x(s), \lambda)\, ds. \tag{4.3}$$

We note that this cumulant can be represented as in equations (2.7.7) and (2.7.13) [**17**, page 175 and 177], where $b_s(x(\cdot)) = H_\lambda(s, x(s), 0)$ and $\widehat{g}_s(\lambda; x(\cdot)) = \widehat{H}(s, x(s), \lambda)$ with $\widehat{H}(s, x, \lambda) = H(s, x, \lambda) - H_\lambda(s, x, 0)\lambda$. We verify for these functions conditions I and II [**17**, pages 215 and 216]. Condition I [**17**, page 215] holds by the form of $H_\lambda(t, x, 0)$ given in Lemma 2.1. By Lemma 2.1, the function $\widehat{H}(s, x, \lambda)$ is nonnegative and convex in λ, is continuous in $(x, \lambda) \in \mathbf{R} \times \mathbf{R}$, and, for all $B \in \mathbf{R}_+$ and $t \in \mathbf{R}_+$,

$$\sup_{|x|+|\lambda| \le B} \sup_{s \le t} \widehat{H}(s, x, \lambda) < \infty,$$

$$\lim_{\lambda \to 0} \sup_{|x| \le B} \sup_{s \le t} \widehat{H}(s, x, \lambda) = 0.$$

This implies condition II [**17**, page 216].

Let Π represent a deviability which solves the maxingale problem associated with $(G_t(\lambda; x(\cdot))$. We need to show that $\Pi(x(\cdot)) = \mathbf{\Pi}(x(\cdot))$. It is always the case that $\Pi(x(\cdot)) \le \mathbf{\Pi}(x(\cdot))$, see [**17**, page 174, Lemma 2.7.11] or [**17**, page 212, (2.8.11)], so we may assume that $\mathbf{\Pi}(x(\cdot)) > 0$. According to the definition of $\mathbf{I}$, if $\mathbf{\Pi}(x(\cdot)) > 0$ then $x(\cdot)$ is absolutely continuous, $x(0) = 0$, $x(t) \ge 0$, and $\dot{x}(t) \le 1$ almost everywhere. Given $x(\cdot)$ with $\mathbf{\Pi}(x(\cdot)) > 0$, we define a function $\lambda(\cdot)$ by setting $\lambda(t) = \widehat{\lambda}(t, x(t), \dot{x}(t))$ if $t > 0$, $x(t) > 0$ and $\dot{x}(t) < 1$. We also let $\lambda(t) = \infty$ if $t > 0$ and $\dot{x}(t) = 1$, and we let $\lambda(t) = -\infty$ if $t > 0$ and $x(t) = 0$. A final specification is that $\lambda(0) = 0$. Since $\widehat{\lambda}(\cdot)$ is continuous on $(0, \infty) \times (0, \infty) \times (-\infty, 1)$ by Lemma 2.3, the function $\lambda(\cdot)$ is a Borel measurable mapping from $\mathbf{R}_+$ into $\mathbf{R} \cup \{\infty\} \cup \{-\infty\}$ if the latter is considered as a compact metric space.

We will first show that $\Pi(x(\cdot)) = \mathbf{\Pi}(x(\cdot))$ for the elements of the set D of $x(\cdot)$ such that $\mathbf{\Pi}(x(\cdot)) > 0$ and the function $\lambda(\cdot)$ is $\mathbf{R}$-valued and is locally bounded in $\mathbf{R}$. By Lemma 2.1, for such $x(\cdot)$, $\dot{x}(t) = H_\lambda(t, x(t), \lambda(t))$ almost everywhere and $x(\cdot)$ is uniquely specified by this differential equation, the function $\lambda(\cdot)$ and the initial condition $x(0) = 0$. Suppose $x(\cdot) \in D$. The required uniqueness is established by checking the hypotheses of Theorem 2.8.14 [**17**, page 213]. More specifically, conditions a) and b) of the theorem need to be checked. Since the function $\lambda(\cdot)$ is locally bounded, it belongs to class $\widehat{\Lambda}$ of Definition 2.8.17 [**17**, page 216]. By Theorem 2.8.19 [**17**, page 217] it satisfies condition a) of Theorem 2.8.14. Since the differential equation $\dot{x}(t) = H_\lambda(t, x(t), \lambda(t))$ almost everywhere has a unique solution with the initial condition $x(0) = 0$, Lemma 2.8.20 part 2 [**17**, page 218] shows that Theorem 2.8.14 condition b) is satisfied as well. According to that theorem, $\Pi(x(\cdot)) = \mathbf{\Pi}(x(\cdot))$. Moreover, $\sup_{y(\cdot)\in p_t^{-1}(p_t x(\cdot))} \Pi(y(\cdot)) = \mathbf{\Pi}_t(x(\cdot))$, where $p_t(x(\cdot)) = (x(t \wedge s), s \in \mathbf{R}_+)$ and $\mathbf{\Pi}_t(x(\cdot)) = \exp(-\int_0^t L(s, x(s), \dot{x}(s))\,ds)$.

By Lemma 2.8.26 and Definition 2.8.24 [**17**, page 222], in order to prove that $\Pi(x(\cdot)) = \mathbf{\Pi}(x(\cdot))$ for arbitrary $x(\cdot)$ with $\mathbf{\Pi}(x(\cdot)) > 0$ it suffices to show that there exists a sequence of functions $x_k(\cdot) \in D$, where $k \in \mathbf{N}$ such that, for all $T \in \mathbf{R}_+$,

$$\lim_{k\to\infty} \sup_{t\in[0,T]} |x_k(t) - x(t)| = 0 \tag{4.4a}$$

and

$$\lim_{k\to\infty} \int_0^T L(t, x_k(t), \dot{x}_k(t))\,dt = \int_0^T L(t, x(t), \dot{x}(t))\,dt. \tag{4.4b}$$

We establish the required properties in stages. As Lemma 2.2 shows, the function $\lambda(\cdot)$ may become unbounded or not $\mathbf{R}$-valued either because $x(\cdot)$ comes close to zero, or because $\dot{x}(\cdot)$ is close to 1, or because $\dot{x}(\cdot)$ is large negative, or because t approaches zero. We will successively deal with all these complications. Firstly, we will show that, given $x(\cdot)$ such that $\mathbf{\Pi}(x(\cdot)) > 0$, there exists a sequence $x_k(\cdot)$ of functions such that $\mathbf{\Pi}(x_k(\cdot)) > 0$, $x_k(t) > \varepsilon_k t$ for all $t > 0$ and some $\varepsilon_k > 0$, and (4.4a) and (4.4b) hold. Then we will show that, given $x(\cdot)$ such that $\mathbf{\Pi}(x(\cdot)) > 0$ and $x(t) > \varepsilon t$ for all $t > 0$ and some $\varepsilon > 0$, there exist $x_k(\cdot)$ such that $\mathbf{\Pi}(x_k(\cdot)) > 0$, $x_k(t) > \varepsilon t$ for $t > 0$, the associated functions $\lambda(\cdot)$ are bounded in neighborhoods of $t = 0$, and (4.4a) and (4.4b) hold. As a final step, we approximate in the sense of (4.4a) and (4.4b) an arbitrary $x(\cdot)$ such that $\mathbf{\Pi}(x(\cdot)) > 0$, $x(t) > \varepsilon t$ for $t > 0$, and $\lambda(\cdot)$ is bounded in a neighbourhood of $t = 0$ with $x_k(\cdot) \in D$.

Here is the first step. Given $x(\cdot)$ such that $\mathbf{\Pi}(x(\cdot)) > 0$, we define $x_k(t) = x(t) \vee (t/k)$. We have that $\dot{x}_k(t) = \dot{x}(t)\mathbf{1}_{\{x(t)\geq t/k\}}(t) + (1/k)\mathbf{1}_{\{x(t)<t/k\}}(t)$ almost everywhere. It is clear that the $x_k(\cdot)$ converge to $x(\cdot)$ for the compact open topology as $k \to \infty$, i.e., (4.4a) holds. We write

$$\int_0^T L(t, x_k(t), \dot{x}_k(t))\,dt = \int_0^T L(t, x(t), \dot{x}(t))\mathbf{1}_{\{x(t)\geq t/k\}}(t)\,dt + \int_0^T L\left(t, \frac{t}{k}, \frac{1}{k}\right)\mathbf{1}_{\{x(t)<t/k\}}(t)\,dt. \tag{4.5}$$

By Lemma 2,3, for $t > 0$ and $k > 1$,

$$L\left(t, \frac{t}{k}, \frac{1}{k}\right) = \frac{\widehat{\lambda}(t, t/k, 1/k)}{k} - H\left(t, \frac{t}{k}, \widehat{\lambda}(t, t/k, 1/k)\right). \tag{4.6}$$

By the expression for $H_\lambda(t, x, 0)$ in Lemma 2.1 part 2, for all k large enough, uniformly in $t \in (0, T]$, $H_\lambda(t, t/k, 0)$ is arbitrarily close to e^{-ct}, so it is greater than $1/k$, which implies by Lemma 2.2 part 1 that $\widehat{\lambda}(t, t/k, 1/k) < 0$ for all k large enough uniformly in $t \in (0, T]$. By (2.1), uniformly in $t \in (0, T]$, for all k large enough, $H(t, t/k, \widehat{\lambda}(t, t/k, 1/k)) \geq \log(1 - e^{-ct})$, so that by (4.6) $0 \leq L(t, t/k, 1/k) \leq -\log(1 - e^{-ct})$. Since, by Lemma 2.3, $L(t, t/k, 1/k) \to L(t, 0, 0)$ as $k \to \infty$ for $t > 0$, on applying Lebesgue's bounded convergence theorem,

$$\lim_{k\to\infty} \int_0^T L\left(t, \frac{t}{k}, \frac{1}{k}\right)\mathbf{1}_{\{x(t)<t/k\}}(t)\,dt = \int_0^T L(t, 0, 0)\mathbf{1}_{\{x(t)=0\}}(t)\,dt.$$

We conclude by (4.5) and monotone convergence that (4.4b) holds.

We proceed with step 2 of the approximation procedure and consider $x(\cdot)$ such that $\mathbf{\Pi}(x(\cdot)) > 0$ and $x(t) > \varepsilon t$ for all $t \in (0, T]$ and some $\varepsilon \in (0, 1)$. We approximate it with a sequence of $x_k(\cdot)$ which have all these properties and, in addition, are such that the associated functions $\lambda(\cdot)$ are bounded in neighbourhoods of $t = 0$. We define $x_k(t)$ for $t \in [0, 1/k]$ as in Lemma 3.1 with $t_1 = 0$, $t_2 = 1/k$ and $a = x(1/k)$, i.e.,

$$x_k(t) = (1 + \alpha_k(e^{-c\alpha_k t} - 1))t\mathbf{1}_{[0,1/k]}(t) + x(t)\mathbf{1}_{(1/k,\infty)}(t),$$

where α_k is the unique nonnegative solution of the equation $(1 + \alpha_k(e^{-c\alpha_k/k} - 1))/k = x(1/k)$. The function $\alpha_k(e^{-c\alpha_k t} - 1)$ being decreasing in t, we have that, for $t \in [0, 1/k]$,

$$(1 + \alpha_k(e^{-c\alpha_k t} - 1))t \geq (1 + \alpha_k(e^{-c\alpha_k/k} - 1))t = kx(1/k)t \geq \varepsilon t,$$

so $x_k(t) \geq \varepsilon t$. It is readily seen that $x_k(t) \leq t$. By Lemma 3.1, $\widehat{\lambda}(t, x_k(t), \dot{x}_k(t)) = \log(1 + (1 - \alpha_k)e^{\alpha_k ct}/\alpha_k)$ for $t \in [0, 1/k]$, so the functions $\lambda_k(t) = \widehat{\lambda}(t, x_k(t), \dot{x}_k(t))$ are bounded on $[0, 1/k]$.

Convergence (4.4a) obviously holds. In order to establish (4.4b), one needs to prove that

$$\lim_{k\to\infty} \int_0^{1/k} L(t, x(t), \dot{x}(t))\, dt = 0, \quad \lim_{k\to\infty} \int_0^{1/k} L(t, x_k(t), \dot{x}_k(t))\, dt = 0.$$

The first limit follows because $\mathbf{I}(x(\cdot)) < \infty$. The second limit follows from the first because by Remark 3.1, $\int_0^{1/k} L(t, x_k(t), \dot{x}_k(t))\, dt \leq \int_0^{1/k} L(t, x(t), \dot{x}(t))\, dt$.

We implement the third step. Suppose that $x(\cdot)$ is such that $\mathbf{\Pi}(x(\cdot)) > 0$, $x(t) > \varepsilon t$ for $t \in (0, T]$, and the associated function $\lambda(\cdot)$ is bounded on $[0, \gamma]$, where $\varepsilon \in (0, 1)$ and $\gamma \in (0, T)$. We define

$$y_k(t) = x(\gamma) + \int_\gamma^t \dot{x}(s) \wedge \left(1 - \frac{1}{k}\right) \mathbf{1}_{\{\dot{x}(s) \geq -k\}}(s)\, ds,$$
$$x_k(t) = x(t)\mathbf{1}_{[0,\gamma]}(t) + y_k(t) \vee (\varepsilon t)\mathbf{1}_{[\gamma,T]}(t).$$

Evidently, $1 - 1/k \geq \dot{x}_k(t) \geq -k$ almost everywhere for $t \in [\gamma, T]$ and all k large enough and $t > x_k(t) \geq \varepsilon t$. By Lemma 2.3, the functions $\lambda_k(t)$ associated with the $x_k(\cdot)$ are bounded on $[\gamma, T]$, so they are bounded on $[0, T]$.

It is readily seen that $x_k(\cdot) \to x(\cdot)$ as $k \to \infty$ for the compact open topology, so we need to establish (4.4b). For $B > 0$ and all $k > B$,

$$
\begin{aligned}
(4.7)\quad & \int_0^T L(t, x_k(t), \dot{x}_k(t))\,dt \\
&= \int_0^{\gamma} L(t, x(t), \dot{x}(t))\,dt \\
&\quad + \int_{\gamma}^{T} L(t, \varepsilon t, \varepsilon)\,\mathbf{1}_{\{y_k(t)<\varepsilon t\}}(t)\,dt \\
&\quad + \int_{\gamma}^{T} L(t, y_k(t), 0)\mathbf{1}_{\{y_k(t)\geq\varepsilon t\}}(t)\mathbf{1}_{\{\dot{x}(t)<-k\}}(t)\,dt \\
&\quad + \int_{\gamma}^{T} L\left(t, y_k(t), \dot{x}(t) \wedge \left(1-\frac{1}{k}\right)\right) \\
&\quad \times \mathbf{1}_{\{y_k(t)\geq\varepsilon t\}}(t)\mathbf{1}_{\{\dot{x}(t)\geq -B\}}(t)\,dt \\
&\quad + \int_{\gamma}^{T} L(t, y_k(t), \dot{x}(t))\mathbf{1}_{\{y_k(t)\geq\varepsilon t\}}(t)\mathbf{1}_{\{-B>\dot{x}(t)\geq -k\}}(t)\,dt.
\end{aligned}
$$

We consider the terms on the righthand side in order. The convergence of the $y_k(\cdot)$ to $x(\cdot)$ implies that $\mathbf{1}_{\{y_k(t)<\varepsilon t\}}(t) \to 0$ as $k \to \infty$. By Lemma 2.3, the function $L(t, x, u)$ is bounded on the set $[\gamma, T] \times [\epsilon\gamma, T] \times [0, 1]$. By Lebesgue's bounded convergence theorem,

$$
(4.8)\qquad \lim_{k\to\infty} \int_{\gamma}^{T} L(t, \varepsilon t, \varepsilon)\,\mathbf{1}_{\{y_k(t)<\varepsilon t\}}(t)\,dt = 0.
$$

Similarly,

$$
(4.9)\qquad \lim_{k\to\infty} \int_{\gamma}^{T} L(t, y_k(t), 0)\mathbf{1}_{\{y_k(t)\geq\varepsilon t\}}(t)\mathbf{1}_{\{\dot{x}(t)<-k\}}(t)\,dt = 0.
$$

Furthermore, since $y_k(t) \to x(t)$ for $t \in [\gamma, T]$ and $L(t, x, u)$ is continuous on $[\gamma, T] \times [\epsilon\gamma, T] \times [-B, 1]$ by Lemma 2.3,

$$
\begin{aligned}
(4.10)\quad \lim_{k\to\infty} \int_{\gamma}^{T} L\Bigg(t, y_k(t), \dot{x}(t) & \\
\wedge \left(1-\frac{1}{k}\right)\Bigg) & \mathbf{1}_{\{y_k(t)\geq\varepsilon t\}}(t)\mathbf{1}_{\{\dot{x}(t)\geq -B\}}(t)\,dt \\
&= \int_{\gamma}^{T} L(t, x(t), \dot{x}(t))\mathbf{1}_{\{\dot{x}(t)\geq -B\}}(t)\,dt.
\end{aligned}
$$

By Lemma 2.3, provided $t > 0$,

$$L(t, y_k(t), \dot{x}(t)) = \widehat{\lambda}(t, y_k(t), \dot{x}(t))\dot{x}(t) - H(t, y_k(t), \widehat{\lambda}(t, y_k(t), \dot{x}(t)))$$

almost everywhere on the set $\{\dot{x}(t) < -B, y_k(t) \geq \varepsilon t\}$. We choose $B > cT$. Since $H_\lambda(t, y_k(t), 0) = e^{-ct} - cy_k(t)$, which is greater than $-B$, we have by Lemma 2.2 part 1 that $\widehat{\lambda}(t, y_k(t), \dot{x}(t)) < 0$ almost everywhere on the set $\{\dot{x}(t) < -B, y_k(t) \geq \varepsilon t\}$. By Lemma 2.2 part 4, $-\widehat{\lambda}(t, y_k(t), \dot{x}(t)) \leq (-\log c - \log(\varepsilon t)) + \log(1 + |\dot{x}(t)|)$ on that set. By (2.1) and (2.2), $H(t, y_k(t), \widehat{\lambda}(t, y_k(t), \dot{x}(t))) \geq \log(1 - e^{-ct}) \wedge (-ct)$. We obtain that almost everywhere on the set $\{\dot{x}(t) < -B, y_k(t) \geq \varepsilon t\}$,

$$\begin{aligned} 0 \leq L(t, y_k(t), \dot{x}(t) &\leq \dot{x}(t)(\log c + \log(\varepsilon t)) \\ &- \dot{x}(t) \log(1 + |\dot{x}(t)|) + (-\log(1 - e^{-ct})) \vee (ct). \end{aligned}$$

By Lemma 2.5, the integral over $[0, T]$ of the righthand side is finite. Since $L(t, y_k(t), \dot{x}(t)) \to L(t, x(t), \dot{x}(t))$ as $k \to \infty$ for $t > 0$ by Lemma 2.3, we derive by Lebesgue's dominated convergence theorem, that

$$\begin{aligned} \lim_{k\to\infty} \int_0^T L(t, y_k(t), \dot{x}(t)) \mathbf{1}_{\{y_k(t) \geq \varepsilon t\}}(t) \mathbf{1}_{\{-k \leq \dot{x}(t) < -B\}}(t)\, dt \\ = \int_0^T L(t, x(t), \dot{x}(t)) \mathbf{1}_{\{\dot{x}(t) < -B\}}(t)\, dt. \end{aligned} \tag{4.11}$$

Putting together (4.7)–(4.11) yields the convergence

$$\lim_{k\to\infty} \int_0^T L(t, x_k(t), \dot{x}_k(t))\, dt = \int_0^T L(t, x(t), \dot{x}(t))\, dt.$$

The uniqueness of a solution to the maxingale problem $(0, G)$ has been proved. By Theorem 5.4.2 [**17**, page 417], the $X^{(n)}(\cdot)$ obey the LDP with $\mathbf{I}$ for the Skorohod J_1-topology on the space of real-valued rightcontinuous functions with lefthand limits defined on the nonnegative halfline. Since $\mathbf{I}(x(\cdot)) = \infty$ if $x(\cdot)$ is discontinuous and the $X^{(n)}(\cdot)$ are random elements of $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$, Theorem 3.2.10 [**17**, page 285] allows us to strengthen the LDP so that it holds for the compact open topology.

We show that $(te^{-ct}, t \in \mathbf{R}_+)$ is the unique zero of $\mathbf{I}$. We have that $\mathbf{I}(x(\cdot)) = 0$ if and only if $\lambda\dot{x}(t) \leq H(t, x(t), \lambda)$ for all $\lambda \in \mathbf{R}$ almost everywhere in t. Since both sides of the inequality coincide for $\lambda = 0$, $\dot{x}(t)$ represents a subgradient of $H(t, x(t), \lambda)$ at $\lambda = 0$. The function $H(t, x, \lambda)$ being convex and differentiable in λ for $t > 0$ implies that $\dot{x}(t) = H_\lambda(t, x(t), 0)$ almost everywhere, i.e., $\dot{x}(t) = e^{-ct} - cx(t)$ almost everywhere, which means that $x(t) = te^{-ct}$. □

Proof of Corollary 1.1. The first assertion follows by Theorem 1.1, the continuous mapping principle, Theorem 3.1 and Remark 3.1. We prove the second part. The argument is an adaptation of the one in the proof of Theorem 3.4 [**5**, page 86]. By the LDP for $X^{(n)}(T)$,

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbf{P}(|X^{(n)}(T) - a| \leq \gamma) \geq -\inf_{x(\cdot):|x(T)-a|<\gamma} \mathbf{I}(x(\cdot))$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}\left(\left\{\sup_{t\in[0,T]} |X^{(n)}(t) - x_{a,T}(t)| \geq \epsilon\right\} \cap \{|X^{(n)}(T) - a| \leq \gamma\}\right) \leq -\inf_{\substack{x(\cdot):|x(T)-a|\leq\gamma \\ \sup_{t\in[0,T]} |x(t)-x_{a,T}(t)|\geq\epsilon}} \mathbf{I}(x(\cdot)).$$

Therefore,

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbf{P}\Big(\sup_{t\in[0,T]} |X^{(n)}(t) - x_a(t)| \geq \epsilon \Big| |X^{(n)}(T) - a| \leq \gamma\Big) \leq -\inf_{\substack{x(\cdot):|x(T)-a|\leq\gamma \\ \sup_{t\in[0,T]} |x(t)-x_{a,T}(t)|\geq\epsilon}} \mathbf{I}(x(\cdot)) + \inf_{x(\cdot):|x(T)-a|<\gamma} \mathbf{I}(x(\cdot)).$$

Since $\mathbf{I}$ is lower compact, the righthand side converges as $\gamma \to 0$ to

$$-\inf_{\substack{x(\cdot):x(T)=a \\ \sup_{t\in[0,T]|x(t)-x_{a,T}(t)|\geq\epsilon}}} \mathbf{I}(x(\cdot)) + \inf_{x(\cdot):x(T)=a} \mathbf{I}(x(\cdot))$$

which is negative by Theorem 3.1 and Remark 3.1. Hence,

$$\lim_{\gamma\to 0} \liminf_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T]} |X^{(n)}(t) - x_{a,T}(t)| \leq \epsilon \middle| |X^{(n)}(T) - a| \leq \gamma\right) = 1. \quad \square$$

5. Proof of Theorem 1.2. Recalling that $\mu^{(n)}$ denotes the measure of jumps of $X^{(n)}(\cdot)$ and $\nu^{(n)}$ denotes its dual predictable projection, we can write

$$X^{(n)}(t) = \int_0^t \int_{\mathbf{R}} x\nu^{(n)}(ds, dx) + M^{(n)}(t), \tag{5.1}$$

where

$$M^{(n)}(t) = \int_0^t \int_{\mathbf{R}} x(\mu^{(n)} - \nu^{(n)})(ds, dx). \tag{5.2}$$

The process $M^{(n)}(\cdot)$ is an $\mathbf{F}^{(n)}$-locally square integrable martingale with predictable quadratic variation process

$$\langle M^{(n)}\rangle(t) = \int_0^t \int_{\mathbf{R}} x^2\nu^{(n)}(ds, dx) - \sum_{0<s\le t} \Bigl(\int_{\mathbf{R}} x\,\nu^{(n)}(\{s\}, dx)\Bigr)^2. \tag{5.3}$$

By (4.1) and (4.2),

$$\int_0^t \int_{\mathbf{R}} x\,\nu^{(n)}(ds, dx) = \frac{1}{n}\sum_{k=1}^{\lfloor nt\rfloor} \Bigl(\Bigl(1 - \frac{c_n}{n}\Bigr)^{k-1} - X^{(n)}\Bigl(\frac{k-1}{n}\Bigr)c_n\Bigr) \tag{5.4a}$$

and

$$\begin{aligned} &\int_0^t \int_{\mathbf{R}} x^2\,\nu^{(n)}(ds, dx) - \sum_{0<s\le t} \Bigl(\int_{\mathbf{R}} x\,\nu^{(n)}(\{s\}, dx)\Bigr)^2 \\ &\quad= \frac{1}{n^2}\sum_{k=1}^{\lfloor nt\rfloor} \Bigl(\Bigl(1 - \frac{c_n}{n}\Bigr)^{k-1} + X^{(n)}\Bigl(\frac{k-1}{n}\Bigr)c_n\Bigl(1 - \frac{c_n}{n}\Bigr) \\ &\qquad - \Bigl(1 - \frac{c_n}{n}\Bigr)^{2(k-1)} + 2\Bigl(1 - \frac{c_n}{n}\Bigr)^{k-1} X^{(n)}\Bigl(\frac{k-1}{n}\Bigr)c_n\Bigr). \end{aligned} \tag{5.4b}$$

According to (5.1), the process $Y^{(n)}(\cdot)$ can be written in the form

$$Y^{(n)}(t) = \sqrt{n}\Bigl(\int_0^t \int_{\mathbf{R}} x\nu^{(n)}(ds, dx) - te^{-ct}\Bigr) + \sqrt{n}M^{(n)}(t).$$

By (5.4a), the equation $te^{-ct} = \int_0^t (e^{-cs} - cse^{-cs})\,ds$, and the convergence of the $X^{(n)}(\cdot)$ to (te^{-ct}) established in Theorem 1.1, for $\epsilon > 0$ and $T > 0$,

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T]} \left|\sqrt{n}\left(\int_0^t \int_{\mathbf{R}} x\nu^{(n)}(ds,dx) - te^{-ct}\right)\right.\right.$$
$$\left.\left. + 2\theta \int_0^t se^{-cs}\,ds + c\int_0^t Y^{(n)}(s)\,ds\right| > \epsilon\right) = 0. \tag{5.5}$$

By (5.3), (5.4b), and the convergence of the $X^{(n)}(\cdot)$ to (te^{-ct}), for $\epsilon > 0$ and $t > 0$,

(5.6)
$$\lim_{n\to\infty} \mathbf{P}\left(\left|n\langle M^{(n)}\rangle(t) - \int_0^t (e^{-cs} + cse^{-cs} - e^{-2cs} + 2cse^{-2cs})\,ds\right| > \epsilon\right) = 0.$$

In order to apply Theorem IX.3.48 [**7**, page 553] (see also Lemma IX.4.4 [**7**, page 555]) it remains to check the Lindeberg condition for the predictable measure of jumps of $Y^{(n)}(\cdot)$ that, for $t > 0$, $\epsilon > 0$ and $\gamma > 0$,

$$\lim_{n\to\infty} \mathbf{P}\left(n\int_0^t \int_{\mathbf{R}} x^2 \mathbf{1}_{\{|x|>\gamma/\sqrt{n}\}}(x)\nu^{(n)}(ds,dx) > \epsilon\right) = 0. \tag{5.7}$$

Clearly,

(5.8)
$$n\int_0^t \int_{\mathbf{R}} x^2 \mathbf{1}_{\{|x|>\gamma/\sqrt{n}\}}(x)\nu^{(n)}(ds,dx) \le \frac{n^2}{\gamma}\int_0^t \int_{\mathbf{R}} x^4 \nu^{(n)}(ds,dx).$$

By (4.1) and the expression for the fourth moment of a binomial distribution as given in [**11**],

$$\begin{aligned}
&\int_0^t \int_{\mathbf{R}} x^4 \nu^{(n)}(ds,dx)\\
&\quad= \frac{1}{n^4}\sum_{k=1}^{\lfloor nt\rfloor} \left(\left(1 - \frac{c_n}{n}\right)^{k-1} + X^{(n)}\left(\frac{k-1}{n}\right)c_n\right.\\
&\qquad + 7\left(nX^{(n)}\left(\frac{k-1}{n}\right)\right)_{(2)} \frac{c_n^2}{n^2}\\
&\qquad \left. + 6\left(nX^{(n)}\left(\frac{k-1}{n}\right)\right)_{(3)} \frac{c_n^3}{n^3} + \left(nX^{(n)}\left(\frac{k-1}{n}\right)\right)_{(4)}, \frac{c_n^4}{n^4}\right),
\end{aligned}$$

where $m_{(i)} = m(m-1)\cdots(m-i+1)$ is the factorial polynomial of order i. Since $X^{(n)}(s) \le s$,

$$\int_0^t \int_{\mathbf{R}} x^4 \, \nu^{(n)}(ds, dx) \le \frac{t}{n^3}(1 + tc_n + 7t^2c_n^2 + 6t^3c_n^3 + t^4c_n^4).$$

The convergence in (5.7) follows by (5.8).

By Theorem IX.3.48 [**7**, page 553], the $Y^{(n)}(\cdot)$ converge in distribution to $Y(\cdot)$ for the Skorohod J_1-topology. The topology can be strengthened to the compact open topology since Y has continuous paths and the $Y^{(n)}(\cdot)$ are random elements of $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$.

Remark 5.1. One could also invoke the results of Section 3 [**14**, Chapter 8] in order to deduce the needed convergence.

6. Proofs of Theorem 1.3 and Corollary 1.3. Theorem 1.3 can be expressed equivalently in terms of large deviation convergence to idempotent processes as in [**17**].

Theorem 6.1. *Suppose that* $(\sqrt{n}/b_n)(c_n - c) \to \widehat{\theta} \in \mathbf{R}$ *as* $n \to \infty$, *where* $c > 0$, $b_n \to \infty$, *and* $b_n/\sqrt{n} \to 0$. *The processes* $(\widehat{Y}^{(n)}(t), t \in \mathbf{R}_+)$, *where* $\widehat{Y}^{(n)}(t) = (\sqrt{n}/b_n)(X^{(n)}(t) - te^{-ct})$, *large deviation converge in distribution in* $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ *at rate* b_n^2 *to the idempotent process* $(\widehat{Y}(t), t \in \mathbf{R}_+)$ *given by the equation*

$$\begin{aligned}\widehat{Y}(t) = &\int_0^t -2\hat{\theta} s e^{-cs}\, ds - c\int_0^t \widehat{Y}(s)\, ds \\ &+ \int_0^t \sqrt{e^{-cs} + c^2s^2e^{-2cs} + cse^{-cs}}\,\dot{\widehat{W}}(s)\, ds,\end{aligned}$$

where $(\widehat{W}(t), t \in \mathbf{R}_+)$ *is a standard idempotent Wiener process. In particular, the* $(\sqrt{n}/b_n)(X^{(n)}(1) - e^{-c})$ *large deviation converge in distribution in* $\mathbf{R}$ *to the Gaussian idempotent variable* $\widehat{N}(-\widehat{\theta}e^{-c}, c^2e^{-2c}/3 + e^{-c})$.

Proof of Theorem 6.1. The plan of the proof is similar to that for Theorem 1.2 except that the results of [**17**] are used rather than the results of [**7**]. To be more specific, we apply Theorem 5.3.8 [**17**, page 413. Since this technique is not as common, more detail is provided.

To start with, we strengthen the convergence in probability of the $X^{(n)}(\cdot)$ to (te^{-ct}) and show exponential convergence at rate b_n^2, i.e., for $\epsilon > 0$ and $T > 0$,

$$\lim_{n\to\infty} \mathbf{P}\big(\sup_{t\in[0,T]} |X^{(n)}(t) - te^{-ct}| > \epsilon\big)^{1/b_n^2} = 0. \tag{6.1}$$

By the LDP for the $X^{(n)}(\cdot)$ and the uniqueness of the zero of $\mathbf{I}$,

$$\limsup_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T]} |X^{(n)}(t) - te^{-ct}| \geq \epsilon\right)^{1/n} \leq \sup_{x(\cdot):\sup_{t\in[0,T]} |x(t)-te^{-ct}|\geq\epsilon} e^{-\mathbf{I}(x(\cdot))} < 1.$$

Limit (6.1) follows since $b_n^2/n \to 0$ as $n \to \infty$.

According to (5.1), the process $\widehat{Y}^{(n)}(\cdot)$ can be written in the form

$$\widehat{Y}^{(n)}(t) = \frac{\sqrt{n}}{b_n}\left(\int_0^t \int_{\mathbf{R}} x\nu^{(n)}(ds, dx) - te^{-ct}\right) + \frac{\sqrt{n}}{b_n} M^{(n)}(t). \tag{6.2}$$

We prove that, in analogy with (5.5), for $\epsilon > 0$ and $T > 0$,

(6.3)

$$\lim_{n\to\infty} \mathbf{P}\left(\sup_{t\in[0,T]} \left|\frac{\sqrt{n}}{b_n}\left(\int_0^t \int_{\mathbf{R}} x\,\nu^{(n)}(ds, dx) - te^{-ct}\right) + 2\widehat{\theta}\int_0^t se^{-cs} ds + c\int_0^t \widehat{Y}^{(n)}(s)\,ds\right| > \epsilon\right)^{1/b_n^2} = 0.$$

Recalling (5.4a), we can write

$$\begin{aligned}\frac{\sqrt{n}}{b_n}\Big(\int_0^t\int_{\mathbf{R}} x\nu^{(n)}(ds,dx)-te^{-ct}\Big)\\ +2\widehat{\theta}\int_0^t se^{-cs}\,ds+c\int_0^t \widehat{Y}^{(n)}(s)\,ds\\ =\int_0^t\bigg(\frac{\sqrt{n}}{b_n}\bigg(\Big(1-\frac{c_n}{n}\Big)^{\lfloor ns\rfloor}-e^{-cs}\bigg)+\widehat{\theta}se^{-cs}\bigg)\,ds\\ -\bigg(\frac{\sqrt{n}}{b_n}(c_n-c)-\widehat{\theta}\bigg)\int_0^t se^{-cs}\,ds\\ -\frac{\sqrt{n}}{b_n}(c_n-c)\int_0^t (X^{(n)}(s)-se^{-cs})\,ds\\ -\frac{\sqrt{n}}{b_n}\int_{\lfloor nt\rfloor/n}^t\bigg(\Big(1-\frac{c_n}{n}\Big)^{\lfloor ns\rfloor}-c_nX^{(n)}(s)\bigg)\,ds.\end{aligned}$$

The first two terms on the right converge to zero locally uniformly in t because $(\sqrt{n}/b_n)(c_n-c)\to\widehat{\theta}$. By (6.1), for arbitrary $\gamma>0$,

$$\lim_{n\to\infty}\mathbf{P}\bigg(\frac{\sqrt{n}}{b_n}(c_n-c)\int_0^t |X^{(n)}(s)-se^{-cs}|\,ds>\gamma\bigg)^{1/b_n^2}=0,$$

which attends to the third term. For similar reasons,

$$\lim_{n\to\infty}\mathbf{P}\bigg(\sup_{t\in[0,T]}\frac{\sqrt{n}}{b_n}\int_{\lfloor nt\rfloor/n}^t\bigg|\Big(1-\frac{c_n}{n}\Big)^{\lfloor ns\rfloor}-c_nX^{(n)}(s)\bigg|\,ds>\gamma\bigg)^{1/b_n^2}=0.$$

Convergence (6.3) has been proved.

We now prove that in analogy with (5.6), for $\epsilon>0$ and $t>0$,

$$\text{(6.4)}\quad \lim_{n\to\infty}\mathbf{P}\bigg(\Big|n\langle M^{(n)}\rangle(t)-\int_0^t(e^{-cs}+cse^{-cs}-e^{-2cs}+2cse^{-2cs})\,ds\Big|>\epsilon\bigg)^{1/b_n^2}=0.$$

By (5.3) and (5.4b),

$$
\begin{aligned}
n\langle M^{(n)}\rangle(t) - \int_0^t (e^{-cs} + cse^{-cs} - e^{-2cs} + 2cse^{-2cs})\,ds
&= \int_0^{\lfloor nt\rfloor/n} \left(\left(1-\frac{c_n}{n}\right)^{\lfloor ns\rfloor} - e^{-cs}\right) ds \\
&\quad + \int_0^{\lfloor nt\rfloor/n} \left(X^{(n)}(s)c_n\left(1-\frac{c_n}{n}\right) - cse^{-cs}\right) ds \\
&\quad - \int_0^{\lfloor nt\rfloor/n} \left(\left(1-\frac{c_n}{n}\right)^{2\lfloor ns\rfloor} - e^{-2cs}\right) ds \\
&\quad + \int_0^{\lfloor nt\rfloor/n} \left(2\left(1-\frac{c_n}{n}\right)^{\lfloor ns\rfloor} X^{(n)}(s)c_n - 2cse^{-2cs}\right) ds \\
&\quad - \int_{\lfloor nt\rfloor/n}^t (e^{-cs} + cse^{-cs} - e^{-2cs} + 2cse^{-2cs})\,ds.
\end{aligned}
$$

Convergence (6.4) now follows by (6.1).

In order to apply Theorem 5.3.8 [**17**, page 413], it remains to check condition $(L_e)_{loc}$ on page 412 where $r_\phi = b_n^2$. It follows if, for $t > 0$, $\epsilon > 0$ and $\gamma > 0$,

$$
(6.5)\qquad \lim_{n\to\infty} \frac{1}{b_n^2} \sum_{k=1}^{\lfloor nt\rfloor} \sup_{0\le u\le t} \int_{\mathbf{R}} e^{\gamma b_n\sqrt{n}|x|} \mathbf{1}_{\{b_n\sqrt{n}|x|>\epsilon\}}(x)\widetilde{\nu}_k^{(n)}(u,dx) = 0.
$$

By (4.2),

$$
\begin{aligned}
\int_{\mathbf{R}} e^{\gamma b_n\sqrt{n}|x|} \mathbf{1}_{\{b_n\sqrt{n}|x|>\epsilon\}}(x)\widetilde{\nu}_k^{(n)}(u,dx)
&\le e^{-\epsilon\sqrt{n}/b_n} \int_{\mathbf{R}} e^{\gamma b_n\sqrt{n}|x|} e^{n|x|}\widetilde{\nu}_k^{(n)}(u,dx) \\
&= e^{-\epsilon\sqrt{n}/b_n} \Biggl(e^{\gamma b_n/\sqrt{n}+1}\left(1-\frac{c_n}{n}\right)^{k-1} \\
&\qquad + \left(1-\frac{c_n}{n} + e^{\gamma b_n/\sqrt{n}+1}\frac{c_n}{n}\right)^{\lfloor n(u\vee 0\wedge((k-1)/n))\rfloor} \\
&\qquad - \left(1-\frac{c_n}{n}\right)^{\lfloor n(u\vee 0\wedge((k-1)/n))\rfloor} \Biggr).
\end{aligned}
$$

The limit in (6.5) follows from the convergence $b_n/\sqrt{n} \to 0$. By Theorem 5.3.8 [**17**, page 413], the $\widehat{Y}^{(n)}(\cdot)$ LD converge at rate b_n^2 to $\widehat{Y}(\cdot)$ for the Skorohod topology. This can be strengthened to the LD convergence in $\mathbf{D}_{co}(\mathbf{R}_+, \mathbf{R})$ as in the proof of Theorem 1.1. □

Proof of Corollary 1.2. The proof proceeds similarly to the proof of Corollary 1.1. Since the integrand in the expression for $\widehat{I}$ is a Carathéodory function and is a strictly convex function of $(y(\cdot), \dot{y}(\cdot))$, by Theorem 4.1 [**4**, page 120], the variational problem of minimizing $\widehat{I}(y(\cdot))$ over absolutely continuous $y(\cdot)$ such that $y(0) = 0$ and $y(T) = b$ has a unique solution. By Theorem 4.12 [**4**, page 125], it satisfies the Euler-Lagrange equation. Solving the latter is straightforward. The limit for the conditional probability is established as in the proof of Corollary 1.1. □

Acknowledgments. The author is thankful to the referee for insightful remarks and for pointing out Janson's work [**8, 9**].

REFERENCES

1. A.D. Barbour, M. Karoński and A. Ruciński, *A central limit theorem for decomposable random variables with applications to random graphs*, J. Combin. Theor. **47** (1989), 125–145.

2. B. Bollobás, *Random graphs*, 2nd edition, Cambridge University Press, Cambridge, 2001.

3. L. Cesari, *Optimization–Theory and applications*, Appl. Math. **17** Springer-Verlag, New York, 1983.

4. B. Dacorogna, *Direct methods in the calculus of variations*, Appl. Math. Sci. **78**, Springer, New York, 2008.

5. M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems*, 2nd edition, Springer, New York, 1998.

6. E.N. Gilbert, *Random graphs*, Ann. Math. Stat. **30** (1959), 1141–1144.

7. J. Jacod and A.N. Shiryaev, *Limit theorems for stochastic processes*, Grundl. Math. Wiss. **288** (2003), Springer-Verlag, Berlin.

8. S. Janson, *Orthogonal decompositions and functional limit theorems for random graph statistics*, Mem. Amer. Math. Soc. **111** (1994), vi+78.

9. ———, *The minimal spanning tree in a complete graph and a functional limit theorem for trees in a random graph*, Random Struct. Algor. **7** (1995), 337–355.

10. S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley, New York, 2000.

11. A.H. Joarder and M. Mahmood, *An inductive derivation of Stirling numbers of the second kind and their applications in statistics*, J. Appl. Math. Dec. Sci. **1** (1997), 151–157.

12. V.F. Kolchin, *Random graphs*, Cambridge University Press, Cambridge, 1999.

13. W. Kordecki, *Normal approximation and isolated vertices in random graphs*, in *Random graphs*, Wiley, Chichester, 1990.

14. R.Sh. Liptser and A.N. Shiryayev, *Theory of Martingales*, Kluwer, Dordrecht, 1989.

15. N. O'Connell, *Some large deviation results for sparse random graphs*, Prob. Theor. Rel. Fields **110** (1998), 277–285.

16. B. Pittel, *On tree census and the giant component in sparse random graphs*, Random Struct. Algor. **1** (1990), 311–341.

17. A. Puhalskii, *Large deviations and idempotent probability*, Chapman & Hall, CRC, Boca Raton, 2001.

18. Y. Punkla and N. Chaidee, *Normal approximation of number of isolated vertices in a random graph*, Thai J. Math. **4** (2006), 1–10.

19. J.M. Steele, *Le Cam's inequality and Poisson approximations*, Amer. Math. Month. **101** (1994), 48–54.

20. R. Vinter, *Optimal control*, Syst. Contr.: Found. Appl., Birkhäuser, Boston, MA, 2000.

University of Colorado Denver, Denver, CO 80217 and Institute for Problems in Information Transmission, Moscow, Russia
Email address: anatolii.puhalskii@ucdenver.edu