

RARE EVENT ASYMPTOTICS FOR EXPLORATION PROCESSES FOR RANDOM GRAPHS

BY SHANKAR BHAMIDI^{1,a}, AMARJIT BUDHIRAJA^{1,b}, PAUL DUPUIS^{2,c} AND RUOYU WU^{3,d}

¹*Department of Statistics and Operations Research, University of North Carolina,* ^a*bhamidi@email.unc.edu,*
^b*budhiraj@email.unc.edu*

²*Division of Applied Mathematics, Brown University,* ^c*paul_dupuis@brown.edu*

³*Department of Mathematics, Iowa State University,* ^d*ruoyu@iastate.edu*

Large deviations for random graph models has been a topic of significant recent research activity. Much work in this area is focused on the class of *dense* random graph models (number of edges in the graph scale as n^2 , where n is the number of vertices) where the theory of graphons has emerged as a principal tool in the study of large deviation properties. These tools do not give a good approach to large deviation problems for random graph models in the sparse regime. The aim of this paper is to study an approach for large deviation problems in this regime by establishing large deviation principles (LDP) on suitable path spaces for certain exploration processes of the associated random graph sequence. Exploration processes are an important tool in the study of sparse random graph models and have been used to understand detailed asymptotics of many functionals of sparse random graphs, such as component sizes, surplus, deviations from trees, etc. In the context of rare event asymptotics of interest here, the point of view of exploration process transforms a large deviation analysis of a static random combinatorial structure to the study of a small noise LDP for certain stochastic dynamical systems with jumps.

Our work focuses on one particular class of random graph models, namely the configuration model; however, the general approach of using exploration processes for studying large deviation properties of sparse random graph models has broader applicability. The goal is to study asymptotics of probabilities of nontypical behavior in the large network limit. The first key step for this is to establish a LDP for an exploration process associated with the configuration model. A suitable exploration process here turns out to be an infinite-dimensional Markov process with transition probability rates that diminish to zero in certain parts of the state space. Large deviation properties of such Markovian models is challenging due to poor regularity behavior of the associated local rate functions. Our proof of the LDP relies on a representation of the exploration process in terms of a system of stochastic differential equations driven by Poisson random measures and variational formulas for moments of nonnegative functionals of Poisson random measures. Uniqueness results for certain controlled systems of deterministic equations play a key role in the analysis. Next, using the rate function in the LDP for the exploration process we formulate a calculus of variations problem associated with the asymptotics of component degree distributions. The second key ingredient in our study is a careful analysis of the infinite-dimensional Euler–Lagrange equations associated with this calculus of variations problem. Exact solutions of these systems of nonlinear differential equations are identified which then provide explicit formulas for decay rates of probabilities of nontypical component degree distributions and related quantities.

Received January 2021; revised April 2021.

MSC2020 subject classifications. Primary 60F10, 60C05, 05C80; secondary 90B15.

Key words and phrases. Large deviation principle, random graphs, sparse regime, diminishing rates, Euler–Lagrange equations, calculus of variations problems, configuration model, branching processes, variational representations, Poisson random measures, exploration process, singular dynamics, giant component.

CONTENTS

1. Introduction	1113
1.1. Proof techniques and overview of contributions	1115
1.2. Organization of the paper	1117
1.3. Notation	1118
2. Assumptions and results	1118
2.1. The configuration model and assumptions	1118
2.2. Large deviation asymptotics for component degree distributions	1119
2.3. Edge-exploration algorithm (EEA)	1123
2.4. An equivalent continuous time exploration process	1124
2.4.1. A simple continuous time construction	1124
2.4.2. A continuous time construction via Poissonization	1125
2.5. Rate function	1127
2.6. LDP and LLN for the exploration process	1128
2.6.1. Law of large number limits	1129
3. Representation and weak convergence of controlled processes	1130
4. Laplace upper bound	1138
5. Laplace lower bound	1139
6. Compact sublevel sets	1143
7. Calculus of variations problem	1145
8. Proof of Theorem 2.6	1151
8.1. Completing the proof of Theorem 2.6	1159
9. Proofs of auxiliary lemmas	1159
9.1. Proof of Lemma 7.1	1159
9.2. Proofs of Lemmas 7.2 and 7.3	1163
9.3. Proof of Lemma 7.4	1164
9.4. Proof of Lemma 7.5	1166
9.5. Proof of Lemma 7.6	1168
10. Proof of LLN	1174
Acknowledgments	1176
Funding	1177
References	1177

1. Introduction. Large deviations for random graph models has been a topic of significant recent research activity (see, e.g., [7, 17, 20–22, 24, 41, 42]). Much work in this area is focused on the class of *dense* random graph models (where the degree of a typical vertex scales like n so that number of edges in the graph scale like n^2 , where n is the number of vertices). In this regime, the theory of graphons obtained under dense graph limits [8, 9, 24, 34] has emerged as a key tool in the study of large deviation asymptotics. More recently there has been significant research activity in studying large deviations for subgraph counts in *moderately sparse* Erdős–Rényi random graphs where edge probabilities scale like $1/n^c$, $0 < c < 1$ (in particular, as in the dense case, the average degree still goes to infinity) using the framework of *nonlinear large deviations* [19]. Results include large deviation estimates for the number of regular subgraphs and cliques [4, 28, 35], sharp asymptotics for the upper tails of subgraph counts by developing quantitative versions of regularity and counting lemmas in extremal graph theory [22] (and their extensions to Erdős–Rényi hypergraphs [23]), relaxation of second order smoothness requirements in [19] and applications to the study of spectral properties of the adjacency matrices of such random graphs [2]. For a recent survey of both of the above regimes, see [18].

In contrast to the above body of work, the focus of this work differs in the following two major ways: first the focus in the current work is on *sparse* random graphs where the average degree of a typical vertex is $O(1)$ so that the number of edges in the graph are $O(n)$ as $n \rightarrow \infty$; second, rather than subgraph counts, in this work we are interested in the large

deviation asymptotics of component degree distributions. The goal of this work is to study an approach for large deviation problems in such sparse regimes by establishing large deviation principles for a class of stochastic dynamical systems, known as the *exploration processes*, that play a central role in the study of sparse random graphs. The idea of using stochastic processes to study large deviation problems for static combinatorial objects has been used previously in several works, for example, in [26] for studying urn models, in [42] for studying Erdős-Rényi random graphs, in [21] in the study of preferential attachment model and in [43] for another type of attachment model. Our work focuses on one particular class of random graph models where the exploration process is an infinite-dimensional stochastic process, namely the configuration model; however similar techniques are expected to be useful for other sparse random graph models as well where tractable dynamic constructions via exploration processes are available.

The *configuration model* refers to a sequence of random graphs with a number of vertices approaching infinity and the degree distribution converging to a prespecified probability distribution $\mathbf{p} = \{p_k\}_{k \in \mathbb{N}}$ on the set of nonnegative integers. This random graph model is a basic object in probabilistic combinatorics; see [36] where sufficient conditions for the existence of a large connected component in a configuration model were given, which then lead to these types of random graphs being used as models for various real world systems, see for example, [38] and [44] and references therein for a comprehensive survey of rigorous results on this model (see also [3, 6] where constructions similar to the configuration model were first used to count graphs with a prescribed degree sequence). This model has become one of the standard workhorses in the study of networks in areas such as epidemiology (see e.g., [39] where epidemics on graphs with prescribed degree distribution are considered) and community detection (where the configuration model forms the basis of one of the most well known techniques called modularity optimization [40], [27], Section 6). In such applications, after observing a real world system, the configuration model with the same degree distribution is used as a “baseline” model to compare against the real world system to judge the existence of atypical events. Thus an important question in such random graph models is to estimate probabilities of atypical structural behaviors, particularly when the system size is large.

In this paper, we are interested in probabilities of events $E^{n,\varepsilon}(\mathbf{q})$ associated with the configuration model random graph G_n on n vertices, described as

$$(1.1) \quad E^{n,\varepsilon}(\mathbf{q}) = \{ \text{there exists a component in } G_n \text{ with } m_k \text{ degree } k \text{ vertices, where} \\ m_k \in [n(q_k - \varepsilon), n(q_k + \varepsilon)], k \in \mathbb{N} \},$$

and where $\mathbf{q} = (q_k)_{k \in \mathbb{N}}$ is such that $0 \leq q_k \leq p_k$ for every k . One of our main results (see Theorem 2.6) shows that, under conditions, for large n and small ε

$$(1.2) \quad P\{E^{n,\varepsilon}(\mathbf{q})\} \approx \exp\{-n[H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p})]\},$$

where for a nonnegative sequence $\mathbf{r} = (r_k)_{k \in \mathbb{N}}$,

$$(1.3) \quad H(\mathbf{r}) \doteq \sum_{k=1}^{\infty} r_k \log r_k - \left(\frac{1}{2} \sum_{k=1}^{\infty} kr_k \right) \log \left(\frac{1}{2} \sum_{k=1}^{\infty} kr_k \right).$$

This result in particular gives asymptotics for probabilities of observing a component of a given size (see Remark 2.12) and explicit formulas for rates of decay of probabilities of observing a D -regular component of a given size in G_n (see Corollaries 2.8 and 2.11); see also Conjectures 2.9 and 2.10 on large deviation asymptotics for the size of the largest component in a D -regular graph.

In order to prove Theorem 2.6 we first study a more general and abstract problem of large deviations for a certain class of stochastic dynamical systems in Theorem 2.16. The starting

point is a dynamical construction of the configuration model given through a discrete time infinite-dimensional Markov chain referred to as the *exploration process* (cf. [31, 37]). As the name suggests, the exploration process is constructed by first appropriately selecting a vertex in the graph and then exploring the neighborhood of the chosen vertex until the component of that vertex is exhausted. After this one moves on to another “unexplored” vertex resulting in successive exploration of components of the random graph until the entire graph has been explored. The stochastic process corresponding to one particular coordinate of this infinite-dimensional Markov chain encodes the number of edges in any given component through the length of its excursions away from zero. The remaining coordinates of this Markov chain can be used to read off the number of vertices of a given degree in any given component of the random graph. See Section 2.3 for a precise description of the state space of this Markov chain. The exploration process can be viewed as a small noise stochastic dynamical system in which the transition steps are of size $O(1/n)$ with n denoting the number of vertices in the random graph. A key ingredient in the proof of Theorem 2.6, is a large deviation principle (LDP) for an infinite-dimensional jump-Markov process that can be viewed as a continuous time analogue of the exploration process. This result, given in Theorem 2.16, is our second main result. As other applications of this theorem, we recover a well-known result on the asymptotics of the largest component in the configuration model due to Molloy and Reed [37] and Janson and Luczak [31], and also present a result (whose proof is omitted) on asymptotics of scaled number of components in a configuration model (see Remark 2.17). The rate function in the LDP given in Theorem 2.16 can be used to formulate a calculus of variations problem associated with the event $E^{n,\varepsilon}(\mathbf{q})$ described in (1.1). This problem is at the heart of our analysis and by studying the corresponding infinite-dimensional system of coupled Euler–Lagrange equations we construct an explicit minimizer in this optimization problem (see Lemma 7.6). The cost associated with the minimizer is the exponent on the right side of (1.2) and provides the exact expression for the decay rate for the probability of interest.

1.1. Proof techniques and overview of contributions. In addition to the study of the asymptotics of the configuration model, one of the main motivations for working on these sets of problems was the development of new techniques for handling large deviations for processes with “degeneracies.” We will give an overview of these contributions in this section.

The exploration process associated with the n th random graph (with n vertices) in the configuration model is described as an \mathbb{R}^∞ -valued “small noise” Markov chain $\{X^n(j)\}_{j \in \mathbb{N}_0}$. Under our assumptions, there exists a $N \in \mathbb{N}$ such that for all $j \geq nN$, $X^n(j) = \mathbf{0}$ for all $n \in \mathbb{N}$. In order to study large deviations for such a sequence, one usually considers a sequence of continuous times processes, or equivalently $\mathbb{C}([0, N] : \mathbb{R}^\infty)$ -valued random variables, obtained by a linear interpolation of $\{X^n(j)\}_{j \in \mathbb{N}_0}$ over intervals of length $1/n$. A large deviations analysis of such a sequence in the current setting is challenging due to the “diminishing rates” feature of the transition kernel (see (2.6)) which in turn leads to poor regularity of the associated local rate function. By diminishing rates we mean the property that probabilities of certain transitions, although nonzero, can get arbitrarily close to 0 as the system becomes large. In the model we consider, the system will go through phases where some state transitions have very low probabilities, that are separated by phases of “regular behavior”, many times. In terms of the underlying random graphs the first type of phase corresponds to time periods in the dynamic construction that are close to the completion of exploration of one component and beginning of exploration of a new component. The poor regularity of the local rate function makes standard approximations of the near optimal trajectory that are used in proofs of large deviation principles for such small noise systems hard to implement.

In order to overcome these difficulties we instead consider a different continuous time process associated with the exploration of the configuration model. This continuous time process is obtained by introducing i.i.d. exponential random times before each step in the edge exploration Markov chain. A precise description of this process is given in terms of stochastic differential equations (SDE) driven by a countable collection of Poisson random measures (PRM), where different PRMs are used to describe the different types of transitions (see Section 2.4). Although the coefficients in this SDE are discontinuous functions, their dependence on the state variable is much more tractable than the state dependence in the transition kernel of the discrete time model.

Large deviations for small noise SDE driven by Brownian motions have been studied extensively both in finite and infinite dimensions. An approach based on certain variational representations for moments of nonnegative functionals of Brownian motions and weak convergence methods [10, 12] has been quite effective in studying a broad range of such systems (cf. references in [15]). A similar variational representation for functionals of a Poisson random measure has been obtained in [15]. There have been several recent papers that have used this representation for studying large deviation problems (see, e.g., [11, 14, 16]). This representation is the starting point of the analysis in the current work as well, however, the application of the representation to the setting considered here leads to new challenges. One key challenge that arises in the proof of the large deviations lower bound can be described as follows. The proof of the lower bound based on variational representations and weak convergence methods, for systems driven by Brownian motions, requires establishing unique solvability of controlled deterministic equations of the form

$$(1.4) \quad dx(t) = b(x(t)) dt + \sigma(x(t)) u(t) dt, \quad x(0) = x_0,$$

where $u \in L^2([0, T] : \mathbb{R}^d)$ (space of square integrable functions from $[0, T]$ to \mathbb{R}^d) is a given control. It turns out that the conditions that are typically introduced for the well-posedness of the original small noise stochastic dynamical system of interest (e.g., Lipschitz properties of the coefficients b and σ) are enough to give the wellposedness of (1.4). For example, when the coefficients are Lipschitz, one can use a standard argument based on Gronwall's lemma and an application of the Cauchy–Schwarz inequality to establish the desired uniqueness property. In contrast, when studying systems driven by a PRM one instead needs to establish wellposedness of controlled equations of the form

$$(1.5) \quad x(t) = x(0) + \int_{[0, t] \times S} 1_{[0, g(x(s))]}(y) \varphi(s, y) ds m(dy), \quad 0 \leq t \leq T,$$

where S is a locally compact metric space, m a locally finite measure on S , $g : \mathbb{R} \rightarrow \mathbb{R}_+$ is a measurable map and the control φ is a nonnegative measurable map on $[0, T] \times S$ which satisfies the integrability property

$$\int_{[0, T] \times S} \ell(\varphi(s, y)) ds m(dy) < \infty,$$

where $\ell(x) = x \log x - x + 1$. If φ were uniformly bounded and g sufficiently regular (e.g., Lipschitz) uniqueness follows once more by a standard Gronwall argument. However, in general if g is not Lipschitz or φ is not bounded (both situations arise in the problem considered here; see, e.g., (2.15)–(2.16)) the problem of uniqueness becomes a challenging obstacle. One of the novel contributions of this work is to obtain uniqueness results for equations of the form (1.5) when certain structural properties are satisfied. The setting we need to consider is more complex than the one described above in that there is an infinite collection of coupled equations (one of which corresponds to the Skorokhod problem for one-dimensional reflected trajectories) that describe the controlled system. However, the basic difficulties can already

be seen for the simpler setting in (1.5). Although for a general φ the unique solvability of equations of the form (1.5) may indeed be intractable, the main idea in our approach is to argue that one can perturb the original φ slightly so that the solution $x(\cdot)$ stays the same and moreover this $x(\cdot)$ is the unique solution of the corresponding equation with the perturbed φ . Furthermore the cost difference between the original and perturbed φ is appropriately small. The uniqueness result given in Lemma 5.1 is a key ingredient in the proof of the lower bound given in Section 5. The proof of the upper bound, via the weak convergence based approach to large deviations relies on establishing suitable tightness and limit characterization results for certain controlled versions of the original small noise system. This proof is given in Section 4.

The rate function in the LDP for the exploration process in Theorem 2.16 is given as a variational formula on an infinite-dimensional path space (see (2.13)). Getting useful information from such an abstract formula in general seems hopeless, however, as we show in this work, for the event considered in (1.1), the variational formula can be used to extract much more explicit information. We begin by observing (see (7.1)) that the event $E^{n,\varepsilon}(\mathbf{q})$ of interest can be written explicitly in terms of the exploration process. Using this and the LDP in Theorem 2.16 one can provide an upper bound for the probability of the event in terms of a quantity $I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q}))$ which can be interpreted (see Section 7 for a precise definition) as the minimal cost for certain controlled analogues of the exploration process to move from the state $(0, \mathbf{p})$ to $(0, \mathbf{p} - \mathbf{q})$ in τ units of time, where $\tau = \frac{1}{2} \sum_{k=1}^{\infty} k q_k$ (see Lemmas 8.2 and 8.3). We then show that this deterministic control problem, which can be reformulated as a calculus of variations problem, admits an explicit solution. This solution is given in Construction 1 and its optimality is studied in Lemma 7.6. Using this optimality property, the complementary lower bound for the probability of interest is given in Lemma 8.4. Lemmas 7.6 and 8.4 form the technical heart of the proof of Theorem 2.6 and rely on a detailed and careful analysis of the infinite-dimensional Euler–Lagrange equations associated with the calculus of variations problem.

1.2. Organization of the paper. The paper is organized as follows. In Section 2 we introduce the configuration model, our main assumptions, and our first main result, Theorem 2.6, on asymptotics of probabilities of $E^{n,\varepsilon}(\mathbf{q})$. We record some consequences of these results for D -regular graphs and subgraphs in Corollaries 2.8 and 2.11. Remark 2.12 discusses another application of this result to the study of asymptotics of probabilities of components of a given size. In Section 2.3 we review the edge-exploration algorithm (EEA) from [31, 37] that gives a dynamical construction of the configuration model. For reasons discussed previously, the large deviation analysis of the discrete time EEA presents several technical obstacles and thus in Section 2.4 we introduce a closely related continuous time jump-Markov process (X^n, Y^n) with values in $(\mathbb{R} \times \mathbb{R}_+^\infty) \times \mathbb{R}$ which is mathematically more tractable. Sections 2.5 and 2.6 present our second main result, Theorem 2.16, that gives a large deviation principle for the sequence $(X^n, Y^n)_{n \in \mathbb{N}}$ in a suitable infinite-dimensional path space. In Section 2.6 we also note two side consequences of Theorem 2.16. The first, given in Section 2.6.1 is a law of large numbers (LLN) result that recovers well-known results of Janson and Luczak (2009) on the asymptotics of the largest component in the configuration model. The second, discussed in Remark 2.17, gives a LDP for the scaled number of components in G_n as $n \rightarrow \infty$.

Section 3 presents the variational representation from [15] for functionals of PRM that is the starting point of our proofs. Some tightness and characterization results that are used both in the upper and lower bound proofs are also given in this section. Next, Section 4 gives the proof of the large deviation upper bound whereas the proof of the lower bound is given in Section 5. Finally, Section 6 establishes the compactness of level sets of the function I_T defined in Section 2.6, thus proving that I_T is a rate function. Together, results of Sections 4, 5 and 6 complete the proof of Theorem 2.16.

We next turn to the proof of Theorem 2.6 which is given in Sections 7–9. First in Section 7 we introduce a calculus of variations problem that is central to the proof of Theorem 2.6. We also introduce (see Construction 1) a candidate minimizer in this optimization problem and present several technical results (Lemmas 7.1–7.6) that are needed for the proof of the optimality property of the candidate minimizer. Using results of Section 7 the proof of Theorem 2.6 is completed in Section 8. Finally, Section 9 contains the proofs of technical lemmas from Section 7 whereas Section 10 presents the proof of the LLN results from Section 2.6.1.

1.3. Notation. The following notation will be used. For a Polish space \mathbb{S} , denote the corresponding Borel σ -field by $\mathcal{B}(\mathbb{S})$. Denote by $\mathcal{P}(\mathbb{S})$ (resp. $\mathcal{M}(\mathbb{S})$) the space of probability measures (resp. finite measures) on \mathbb{S} , equipped with the topology of weak convergence. Denote by $\mathbb{C}_b(\mathbb{S})$ (resp. $\mathbb{M}_b(\mathbb{S})$) the space of real bounded and continuous functions (resp. bounded and measurable functions). For $f : \mathbb{S} \rightarrow \mathbb{R}$, let $\|f\|_\infty \doteq \sup_{x \in \mathbb{S}} |f(x)|$. For a Polish space \mathbb{S} and $T > 0$, denote by $\mathbb{C}([0, T] : \mathbb{S})$ (resp. $\mathbb{D}([0, T] : \mathbb{S})$) the space of continuous functions (resp. right continuous functions with left limits) from $[0, T]$ to \mathbb{S} , endowed with the uniform topology (resp. Skorokhod topology). We recall that a collection $\{X^n\}$ of \mathbb{S} -valued random variables on some probability space (Ω, \mathcal{F}, P) is said to be tight, if for each $\varepsilon > 0$ there is a compact set $K \subset \mathbb{S}$ such that $\sup_n P(X^n \in K^c) \leq \varepsilon$. A sequence of $\mathbb{D}([0, T] : \mathbb{S})$ -valued random variables is said to be \mathcal{C} -tight if it is tight in $\mathbb{D}([0, T] : \mathbb{S})$ and every weak limit point takes values in $\mathbb{C}([0, T] : \mathbb{S})$ a.s. We use the symbol ‘ \Rightarrow ’ to denote convergence in distribution.

We denote by \mathbb{R}^∞ the space of all real sequences which is identified with the countable product of copies of \mathbb{R} . This space is equipped with the usual product topology. For $\mathbf{x} = (x_k)_{k \in \mathbb{N}}$, $\mathbf{y} = (y_k)_{k \in \mathbb{N}}$, we write $\mathbf{x} \leq \mathbf{y}$ if $x_k \leq y_k$ for each $k \in \mathbb{N}$. We will use the notation $a \doteq b$ to signify that the definition of a is given by the quantity b . Let $\mathcal{C} \doteq \mathbb{C}([0, T] : \mathbb{R})$, $\mathcal{C}_\infty \doteq \mathbb{C}([0, T] : \mathbb{R}^\infty)$, $\mathcal{D} \doteq \mathbb{D}([0, T] : \mathbb{R})$, $\mathcal{D}_\infty \doteq \mathbb{D}([0, T] : \mathbb{R}^\infty)$. Let $x^+ \doteq \max\{x, 0\}$ for $x \in \mathbb{R}$. Denote by \mathbb{R}_+ the set of all nonnegative real numbers. Let $\mathbb{N}_0 \doteq \mathbb{N} \cup \{0\}$. Cardinality of a set A is denoted by $|A|$. For $n \in \mathbb{N}$, let $[n] \doteq \{1, 2, \dots, n\}$. We use the following conventions: $0 \log 0 = 0$, $0 \log(x/0) = 0$ for $x \geq 0$ and $x \log(x/0) = \infty$ for $x > 0$.

2. Assumptions and results. Fix $n \in \mathbb{N}$. We start by describing the construction of the configuration model of random graphs with vertex set $[n]$. A detailed description and further references for the configuration model can be found in [44], Chapter 7.

2.1. The configuration model and assumptions. Let $\mathbf{d}(n) = \{d_i^{(n)}\}_{i \in [n]}$ be a degree sequence, namely a sequence of nonnegative integers such that $\sum_{i=1}^n d_i^{(n)}$ is even. Let $2m^{(n)} \doteq \sum_{i=1}^n d_i^{(n)}$. We will usually suppress the dependence of $d_i^{(n)}$ and $m^{(n)}$ on n in the notation. Using the sequence $\{d_i\}$ we construct a random graph on n labelled vertices $[n]$ as follows: (i) Associate with each vertex $i \in [n]$ d_i half-edges. (ii) Perform a uniform random matching on the $2m$ half-edges to form m edges so that every edge is composed of two half-edges. This procedure creates a random multigraph $G([n], \mathbf{d}(n))$ with m edges, allowing for multiple edges between two vertices and self-loops, and is called the *configuration model* with degree sequence $\mathbf{d}(n)$. Since we are concerned with connectivity properties of the resulting graph, vertices with degree zero play no role in our analysis, and therefore, we assume that $d_i > 0$ for all $i \in [n]$, $n \geq 1$. We make the following additional assumptions.

ASSUMPTION 2.1. There exists a probability distribution $\mathbf{p} \doteq \{p_k\}_{k \in \mathbb{N}}$ on \mathbb{N} such that, writing $n_k^{(n)} \doteq |\{i \in [n] : d_i = k\}|$ for the number of vertices with degree k , $n_k^{(n)}/n \rightarrow p_k$ as $n \rightarrow \infty$, for all $k \in \mathbb{N}$.

We will also usually suppress the dependence of $n_k^{(n)}$ on n in the notation. We make the following assumption on moments of the degree distribution.

ASSUMPTION 2.2. There exists some $\varepsilon_p \in (0, \infty)$ such that $\sup_{n \in \mathbb{N}} \sum_{k=1}^{\infty} \frac{n_k}{n} k^{1+\varepsilon_p} < \infty$.

The above two assumptions will be made throughout this work.

REMARK 2.3.

(i) Note that Assumptions 2.1 and 2.2, along with Fatou's lemma, imply that $\sum_{k=1}^{\infty} p_k \times k^{1+\varepsilon_p} < \infty$. Conversely, if $\sum_{k=1}^{\infty} p_k k^{\lambda} < \infty$ for some $\lambda \in (4, \infty)$ and $\{D_i\}_{i \in \mathbb{N}}$ is a sequence of i.i.d. \mathbb{N} -valued random variables with common distribution $\{p_k\}_{k \in \mathbb{N}}$, then using a Borel-Cantelli argument it can be shown that for a.e. ω , Assumptions 2.1 and 2.2 are satisfied with $d_i = D_i(\omega)$, $i \in [n]$, $n \in \mathbb{N}$ and $\varepsilon_p = \frac{\lambda}{4} - 1$.

(ii) Under Assumptions 2.1 and 2.2, $\mu \doteq \sum_{k=1}^{\infty} k p_k < \infty$ and the total number of edges $m = \frac{1}{2} \sum_{i=1}^n d_i$ satisfies $\frac{m}{n} \rightarrow \frac{1}{2} \sum_{k=1}^{\infty} k p_k$ as $n \rightarrow \infty$.

2.2. *Large deviation asymptotics for component degree distributions.* We will say that a component of $G([n], \mathbf{d}(n))$ has degree configuration $\{\bar{n}_k\}$ if the component has \bar{n}_k vertices with degree k , for $k \in \mathbb{N}$. Given $\mathbf{0} \leq \mathbf{q} = (q_k, k \in \mathbb{N}) \leq \mathbf{p}$, we are interested in the asymptotic exponential rate of decay of the probability of the event $E^{n,\varepsilon}(\mathbf{q})$ introduced in (1.1) that corresponds to the existence of a component in $G([n], \mathbf{d}(n))$ with degree configuration $\{\bar{n}_k\}$ satisfying $(q_k - \varepsilon)n \leq \bar{n}_k \leq (q_k + \varepsilon)n$, $k \in \mathbb{N}$, namely, we want to characterize $\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} \log P\{E^{n,\varepsilon}(\mathbf{q})\}$. Note that for there to exist a component with degree configuration $\{nq_k\}$ we must have $\sum_{k=1}^{\infty} k q_k \geq 2(\sum_{k=1}^{\infty} q_k - \frac{1}{n})$. We will in fact assume a slightly stronger condition:

$$(2.1) \quad \sum_{k=1}^{\infty} k q_k > 2 \sum_{k=1}^{\infty} q_k.$$

This condition says that there are strictly more edges than vertices in the component. Define $\beta \doteq \beta(\mathbf{q})$ as follows: $\beta = 0$ when $q_1 = 0$, and when $q_1 > 0$, $\beta \in (0, 1)$ is the unique solution (see Remark 2.4 below) of the equation

$$\sum_{k=1}^{\infty} k q_k = (1 - \beta^2) \sum_{k=1}^{\infty} \frac{k q_k}{1 - \beta^k}.$$

Define the function $K(\mathbf{q})$ by

$$(2.2) \quad K(\mathbf{q}) \doteq \left(\frac{1}{2} \sum_{k=1}^{\infty} k q_k \right) \log(1 - \beta(\mathbf{q})^2) - \sum_{k=1}^{\infty} q_k \log(1 - \beta(\mathbf{q})^k)$$

and with $H(\cdot)$ as in (1.3) define

$$(2.3) \quad \tilde{I}_1(\mathbf{q}) \doteq H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}) + K(\mathbf{q}).$$

REMARK 2.4. The existence and uniqueness of $\beta(\mathbf{q})$ can be seen as follows. For $\alpha \in (0, 1)$ consider

$$\alpha F(\alpha) \doteq \sum_{k=1}^{\infty} k q_k - (1 - \alpha^2) \sum_{k=1}^{\infty} \frac{k q_k}{1 - \alpha^k} = \alpha \left(\sum_{k=3}^{\infty} \frac{\alpha - \alpha^{k-1}}{1 - \alpha^k} k q_k - q_1 \right).$$

For $k \geq 3$ and $\alpha \in (0, 1)$ let $F_k(\alpha) \doteq (\alpha - \alpha^{k-1})/(1 - \alpha^k)$. It is easily verified that $F_k(\cdot)$ is strictly increasing on $(0, 1)$. Thus for $\alpha \in (0, 1)$, $0 = F_k(0+) < F_k(\alpha) < F_k(1-) = \frac{k-2}{k}$, and so

$$-q_1 = F(0+) < F(\alpha) < F(1-) = \sum_{k=3}^{\infty} (k-2)q_k - q_1.$$

Since F is continuous on $(0, 1)$, $-q_1 < 0$ and $\sum_{k=3}^{\infty} (k-2)q_k - q_1 = \sum_{k=1}^{\infty} kq_k - 2\sum_{k=1}^{\infty} q_k > 0$, we have the existence and uniqueness of $\beta(\mathbf{q})$.

REMARK 2.5. We note that for every $\mathbf{0} \leq \mathbf{q} = (q_k, k \in \mathbb{N}) \leq \mathbf{p}$, $K(\mathbf{q})$ and $H(\mathbf{q})$ are finite. Indeed, the finiteness of $K(\mathbf{q})$ is immediate from Assumption 2.2. To see the finiteness of $H(\mathbf{q})$, note that on the one hand $\sum_{k=1}^{\infty} q_k \log q_k \leq 0$ while on the other hand

$$\begin{aligned} \sum_{k=1}^{\infty} q_k \log q_k &= \sum_{k=1}^{\infty} q_k \log \frac{q_k}{2^{-(k+1)}} - (\log 2) \sum_{k=1}^{\infty} (k+1)q_k \\ &\geq -\left(1 - \sum_{k=1}^{\infty} q_k\right) \log \frac{1 - \sum_{k=1}^{\infty} q_k}{2^{-1}} - (\log 2) \sum_{k=1}^{\infty} (k+1)q_k > -\infty, \end{aligned}$$

where the first inequality follows from nonnegativity of relative entropy and putting mass $1 - \sum_{k=1}^{\infty} q_k$ on $k = 0$, and the last inequality once more uses Assumption 2.2.

The following result gives asymptotics of the event $E^{n,\varepsilon}(\mathbf{q})$. The proof of the theorem, which is based on Theorem 2.16, is given in Section 8.

THEOREM 2.6. *Suppose $\mathbf{0} \leq \mathbf{q} \leq \mathbf{p}$ and that (2.1) is satisfied. Then*

(i) *(Upper bound) when $p_1 = 0$, we have $\beta(\mathbf{q}) = 0$, $K(\mathbf{q}) = 0$ and*

$$\limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P\{E^{n,\varepsilon}(\mathbf{q})\} \leq -\tilde{I}_1(\mathbf{q}).$$

(ii) *(Lower bound)*

$$\liminf_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log P\{E^{n,\varepsilon}(\mathbf{q})\} \geq -\tilde{I}_1(\mathbf{q}).$$

REMARK 2.7. The proof of Theorem 2.6 relies on a large deviation principle for a certain exploration process (see Section 2.4) that is given in Theorem 2.16. The latter result does not require the condition $p_1 = 0$. Also note that the lower bound in Theorem 2.6 does not require the condition $p_1 = 0$ either. One can also give an upper bound (without requiring $p_1 = 0$) in terms of a variational formula given by the right side of (8.4). When $p_1 = 0$, this variational expression can be simplified and is seen to be equal to $-\tilde{I}_1(\mathbf{q})$. This is shown in Lemma 8.3 whose proof crucially relies on the property $p_1 = 0$. Whether the two expressions are equal in general when $p_1 \neq 0$ remains an open problem.

As an immediate corollary of Theorem 2.6 we have the following result for D -regular graphs, that is, graphs such that each vertex is of degree D . In the following $\lim \star$ represents either \limsup or \liminf .

COROLLARY 2.8 (D -regular graphs). *Suppose that there exists some $D \in \mathbb{N}$ with $D \geq 3$, such that $p_k = 0$, $n_k = 0$ for $k \neq D$ and $p_D = 1$, $n_D = n$. Fix $q_D \in (0, 1]$ and denote by*

$E_D^{n,\varepsilon}(\mathbf{q})$ the event that there is a component of size $N_D \in [n(q_D - \varepsilon), n(q_D + \varepsilon)]$. Then

$$(2.4) \quad \begin{aligned} & \lim_{\star \varepsilon \rightarrow 0} \lim_{\star n \rightarrow \infty} \frac{1}{n} \log P\{E_D^{n,\varepsilon}(\mathbf{q})\} \\ &= \left(1 - \frac{D}{2}\right)(q_D \log q_D + (1 - q_D) \log(1 - q_D)). \end{aligned}$$

PROOF. Let $q_k = 0$ for $k \in \mathbb{N} \setminus \{D\}$ and let $\mathbf{q} = \{q_k, k \in \mathbb{N}\}$. Then since $p_1 = 0$, we have $\beta(\mathbf{q}) = 0$ and $K(\mathbf{q}) = 0$. Using (2.3) we have

$$\begin{aligned} \tilde{I}_1(\mathbf{q}) &= H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}) + K(\mathbf{q}) \\ &= q_D \log q_D - \frac{D q_D}{2} \log\left(\frac{D q_D}{2}\right) + (1 - q_D) \log(1 - q_D) \\ &\quad - \frac{D - D q_D}{2} \log\left(\frac{D - D q_D}{2}\right) + \frac{D}{2} \log\left(\frac{D}{2}\right) \\ &= \left(1 - \frac{D}{2}\right)(q_D \log q_D + (1 - q_D) \log(1 - q_D)). \end{aligned}$$

The result then follows from Theorem 2.6. \square

We note that the expression (2.4) has the same form when q_D is replaced by $1 - q_D$. This suggests that the most likely way of having a component of size around nq_D in D -regular graphs is to let almost all of the remaining $n(1 - q_D)$ vertices be in one component. Indeed, conditioning on having a component of size around nq_D , the remaining vertices can be viewed as a smaller configuration model of D -regular graphs with about $n(1 - q_D)$ vertices. It then follows from the well known results for the asymptotics of the largest component in the configuration model [31, 37] (and Theorem 2.20) that these remaining vertices are in one component with high probability.

Based on these observations we make the following conjecture.

CONJECTURE 2.9 (D -regular graphs, multiple components). *Suppose that there exists some $D \in \mathbb{N}$ with $D \geq 3$, such that $p_k = 0$, $n_k = 0$ for $k \neq D$ and $p_D = 1$, $n_D = n$. Fix $M \in \mathbb{N}$ and $q_D^{(i)} \in (0, 1]$ for each $i = 1, \dots, M$, such that $\sum_{i=1}^M q_D^{(i)} \leq 1$. Let $q_k^{(i)} = 0$ for $k \in \mathbb{N} \setminus \{D\}$ and let $\mathbf{q}^{(i)} = \{q_k^{(i)}, k \in \mathbb{N}\}$, for each $i = 1, \dots, M$. Let $\mathbf{q}^{(M+1)} = \mathbf{p} - \sum_{i=1}^M \mathbf{q}^{(i)}$. Denote by $E_D^{n,\varepsilon,M}$ the event that there are components of sizes $N_D^{(i)} \in [n(q_D^{(i)} - \varepsilon), n(q_D^{(i)} + \varepsilon)]$, $i = 1, \dots, M$. Then*

$$\lim_{\star \varepsilon \rightarrow 0} \lim_{\star n \rightarrow \infty} \frac{1}{n} \log P\{E_D^{n,\varepsilon,M}\} = \sum_{i=1}^{M+1} H(\mathbf{q}^{(i)}) - H(\mathbf{p}) = \left(1 - \frac{D}{2}\right) \sum_{i=1}^{M+1} q_D^{(i)} \log q_D^{(i)}.$$

We also note that for each fixed $a \in [0, 1]$, the function $[0, a] \ni x \mapsto x \log x + (a - x) \log(a - x) \in (-\infty, 0]$ is maximized at $x = 0$ and $x = a$. This suggests that, the most likely way for the largest component to be of certain size, is to let as many of the remaining components as possible have such a size. Based on this we make the following conjecture on the large deviation behavior of the largest component size for D -regular graphs.

CONJECTURE 2.10 (D -regular graphs, largest component). *Suppose that there exists some $D \in \mathbb{N}$ with $D \geq 3$, such that $p_k = 0$, $n_k = 0$ for $k \neq D$ and $p_D = 1$, $n_D = n$. For each $x \in [0, 1]$, let $q_D^{(x)} = x$, $q_k^{(x)} = 0$ for $k \in \mathbb{N} \setminus \{D\}$, and $\mathbf{q}^{(x)} = \{q_k^{(x)}, k \in \mathbb{N}\}$. Denote by M^n the*

size of the largest component. Then $\frac{M^n}{n}$ satisfies a large deviation principle in \mathbb{R}_+ with rate function I_{\max} defined by

$$\begin{aligned} I_{\max}(x) &= k(x)H(\mathbf{q}^{(x)}) + H(\mathbf{q}^{(1-xk(x))}) - H(\mathbf{p}) \\ &= \left(1 - \frac{D}{2}\right)(xk(x)\log x + (1 - xk(x))\log(1 - xk(x))) \end{aligned}$$

for $x \in [0, 1]$ and $I_{\max}(x) = \infty$ otherwise, where $k(x) = \lfloor \frac{1}{x} \rfloor$ is the largest integer such that $xk(x) \leq 1$.

Recall that $\mu \doteq \sum_{k=1}^{\infty} kp_k < \infty$. The following result gives bounds on probabilities of observing a D -regular subgraph in a configuration model with a general degree sequence (p_k) .

COROLLARY 2.11. *Suppose that Assumptions 2.1 and 2.2 hold. Also suppose that $p_1 = 0$. Fix $D \in \mathbb{N}$ with $D \geq 3$ such that $p_D > 0$. Fix $q_D \in (0, p_D]$. Denote by $E^{n,\varepsilon}(\mathbf{q})$ the event that the graph has a component that is D -regular and has size $N_D \in [n(q_D - \varepsilon), n(q_D + \varepsilon)]$. Then*

$$\begin{aligned} &\lim_{\star \varepsilon \rightarrow 0} \lim_{\star n \rightarrow \infty} \frac{1}{n} \log P\{E^{n,\varepsilon}(\mathbf{q})\} \\ &= (q_D \log q_D + (p_D - q_D) \log(p_D - q_D) - p_D \log p_D) \\ &\quad - \left(\frac{Dq_D}{2} \log\left(\frac{Dq_D}{2}\right) + \frac{\mu - Dq_D}{2} \log\left(\frac{\mu - Dq_D}{2}\right) - \frac{\mu}{2} \log\left(\frac{\mu}{2}\right) \right). \end{aligned}$$

PROOF. Let $q_k = 0$ for $k \in \mathbb{N} \setminus \{D\}$ and let $\mathbf{q} = (q_k, k \in \mathbb{N})$. As before, since $q_1 = 0$, we have $\beta(\mathbf{q}) = 0$ and $K(\mathbf{q}) = 0$. Using (2.3) we have

$$\begin{aligned} \tilde{I}_1(\mathbf{q}) &= H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}) + K(\mathbf{q}) \\ &= q_D \log q_D - \frac{Dq_D}{2} \log\left(\frac{Dq_D}{2}\right) \\ &\quad + (p_D - q_D) \log(p_D - q_D) + \sum_{k \neq D} p_k \log p_k - \frac{\mu - Dq_D}{2} \log\left(\frac{\mu - Dq_D}{2}\right) \\ &\quad - \sum_{k=1}^{\infty} p_k \log p_k + \frac{\mu}{2} \log\left(\frac{\mu}{2}\right) \\ &= (q_D \log q_D + (p_D - q_D) \log(p_D - q_D) - p_D \log p_D) \\ &\quad - \left(\frac{Dq_D}{2} \log\left(\frac{Dq_D}{2}\right) + \frac{\mu - Dq_D}{2} \log\left(\frac{\mu - Dq_D}{2}\right) - \frac{\mu}{2} \log\left(\frac{\mu}{2}\right) \right). \end{aligned}$$

The result then follows from Theorem 2.6. \square

REMARK 2.12. Theorem 2.6 can be used to extract other asymptotic results. We give below one example without proof. Suppose that Assumptions 2.1 and 2.2 hold. Also suppose that $p_1 = p_2 = 0$. Let $r \in (0, 1]$ and denote by $E_r^{n,\varepsilon}$ the event that the graph has a component that has size $N_r \in [n(r - \varepsilon), n(r + \varepsilon)]$. Then

$$\lim_{\star \varepsilon \rightarrow 0} \lim_{\star n \rightarrow \infty} \frac{1}{n} \log P\{E_r^{n,\varepsilon}\} = \inf_{0 \leq \mathbf{q} \leq \mathbf{p}: \mathbf{q} \cdot \mathbf{1} = r} \{H(\mathbf{p}) - H(\mathbf{q}) - H(\mathbf{p} - \mathbf{q})\}.$$

REMARK 2.13. There is an important connection between the configuration model and the uniform distribution on the space of all simple graphs (namely graphs which have no multiple edges and self-loops) with a prescribed degree distribution which we now describe. Given a degree sequence $\mathbf{d}(n)$, let $\mathbb{G}([n], \mathbf{d}(n))$ be the set of all (simple) graphs on vertex set $[n]$ with degree sequence $\mathbf{d}(n)$. Let $\mathbb{UM}_n(\mathbf{d}(n))$ denote the uniform measure on $\mathbb{G}([n], \mathbf{d}(n))$. Then as is well known (see, e.g., [44], Proposition 7.15), the configuration model satisfies the property that the conditional distribution of $G([n], \mathbf{d}(n))$, given the event that $G([n], \mathbf{d}(n))$ is simple, is $\mathbb{UM}_n(\mathbf{d}(n))$. Further, by [30], under the assumptions of the current paper $P(G([n], \mathbf{d}(n)) \text{ is simple}) \rightarrow e^{-(v/2+v^2/4)}$ where $v = \sum_k k(k-1)p_k / \sum_k kp_k$. These observations suggest a natural approach to asymptotic questions of the form studied in the current work for (simple) graphs with a prescribed degree distribution. In particular by an elementary Bayes formula calculation it follows that if

$$(2.5) \quad \frac{\log P(G([n], \mathbf{d}(n)) \text{ is simple} \mid E^{n,\varepsilon}(\mathbf{q}))}{n} \rightarrow 0,$$

then Theorem 2.6 will continue to hold with the configuration model replaced with the uniform distribution on the space of simple graphs with prescribed degree sequence. In general, characterizing the asymptotics of quantities as in (2.5) is key to the large deviation analysis of $\mathbb{UM}_n(\mathbf{d}(n))$. Study of these questions is deferred to future work.

2.3. *Edge-exploration algorithm (EEA).* Given a degree sequence $\mathbf{d}(n)$, we now describe a well-known dynamic construction of the configuration model $G([n], \mathbf{d}(n))$ given in [31, 37] by sequentially matching half-edges. Tracking functionals of this dynamic construction, in particular hitting times of zero of the number of so-called active edges (see below) reveals component size information of $G([n], \mathbf{d}(n))$. Construction given below closely follows [31]. This algorithm traverses the graph by exploring all its edges, unlike typical graph exploration algorithms, which sequentially explore vertices. At each stage of the algorithm, every vertex in $[n]$ is in one of two possible states, sleeping or awake, while each half-edge is in one of three states: sleeping (unexplored), active or dead (removed). The exploration process sequentially visits vertices, *awakening* vertices while *activating* or *killing* half-edges.

Write $\mathcal{S}_{\mathbb{V}}(j)$ for the set of sleeping vertices at step j and similarly let $\mathcal{S}_{\mathbb{E}}(j)$, $\mathcal{A}_{\mathbb{E}}(j)$ be the set of sleeping and active half-edge at step j . We call a half-edge “living” if it is either sleeping or active. Initialize by setting all vertices and half-edges to be in the sleeping state. For step $j \geq 0$, write $A(j) \doteq |\mathcal{A}_{\mathbb{E}}(j)|$ for the number of active half-edges and $V_k(j)$ for the number of sleeping vertices $v \in \mathcal{S}_{\mathbb{V}}(j)$ with degree k . Write $V(j) \doteq (V_k(j), k \in \mathbb{N})$ for the corresponding vector in \mathbb{R}_+^∞ . At step $j = 0$, all vertices and half-edges are asleep hence $A(0) = 0$ and $V_k(0) = n_k$ for $k \geq 1$. The exploration process proceeds as follows:

- (1) If the number of active half-edges and sleeping vertices is zero, that is, $A(j) = 0$ and $V(j) = \mathbf{0}$, all vertices and half-edges have been explored and we terminate the algorithm.
- (2) If $A(j) = 0$ and $V(j) \neq \mathbf{0}$, so there exist sleeping vertices, pick one such vertex with probability proportional to its degree and mark the vertex as awake and all its half-edges as active. Thus the transition $(A(j), V(j))$ to $(A(j+1), V(j+1))$ at step $j+1$ takes the form

$$(0, v) \mapsto (k, v - e_k) \text{ with probability } \frac{k v_k}{\sum_{i=1}^{\infty} i v_i}, \quad k \in \mathbb{N},$$

where e_k is the k th unit vector.

- (3) If $A(j) > 0$, pick an active half-edge uniformly at random, pair it with another uniformly chosen living half-edge (either active or sleeping), say e^* , merge both half-edges to form a full edge and kill both half-edges. If e^* was sleeping when picked, wake the vertex

corresponding to the half-edge e^* , and mark all its other half-edges active. Thus in this case the transition takes the form

$$(a, \mathbf{v}) \mapsto (a - 2, \mathbf{v}) \text{ with probability } \frac{a - 1}{\sum_{i=1}^{\infty} i v_i + a - 1},$$

$$(a, \mathbf{v}) \mapsto (a + k - 2, \mathbf{v} - e_k) \text{ with probability } \frac{k v_k}{\sum_{i=1}^{\infty} i v_i + a - 1}, \quad k \in \mathbb{N}.$$

The statements in (2) and (3) can be combined as follows: If $A(j) \neq 0$ or $\mathbf{V}(j) \neq \mathbf{0}$, then the transition $(A(j), \mathbf{V}(j))$ to $(A(j+1), \mathbf{V}(j+1))$ takes the form

$$(a, \mathbf{v}) \mapsto (a - 2 \cdot 1_{\{a>0\}}, \mathbf{v}) \text{ with probability } \frac{(a - 1)^+}{\sum_{i=1}^{\infty} i v_i + (a - 1)^+},$$

$$(2.6) \quad (a, \mathbf{v}) \mapsto (a + k - 2 \cdot 1_{\{a>0\}}, \mathbf{v} - e_k) \text{ with probability } \frac{k v_k}{\sum_{i=1}^{\infty} i v_i + (a - 1)^+},$$

$$k \in \mathbb{N}.$$

The random graph $G([n], \mathbf{d}(n))$ formed at the termination of the above algorithm has the same distribution as the configuration model with degree sequence $\mathbf{d}(n)$ [31, 36].

REMARK 2.14. We note that for $j > 0$, $A(j) = 0$ if and only if the exploration of a component in the random graph $G([n], \mathbf{d}(n))$ is completed at step j . Thus the number of edges in a component equals the length of an excursion of $\{A(j)\}$ away from 0 and the largest excursion length gives the size of the largest component, namely the number of edges in the component with maximal number of edges. The vertices in each component are those that are awakened during corresponding excursions.

Note that at each step in the EEA, either a new vertex is woken up or two half-edges are killed. Since there are a total of n vertices and $2m$ half-edges, we have from Assumptions 2.1 and 2.2 that the algorithm terminates in at most $m + n \leq nL$ steps where $L = 1 + \lfloor \sup_n \frac{1}{2} \sum_{k=1}^{\infty} k \frac{n_k}{n} \rfloor < \infty$. We define $A(j) \equiv 0$ and $\mathbf{V}(j) \equiv \mathbf{0}$ for all $j \geq j_0$ where j_0 is the step at which the algorithm terminates.

2.4. An equivalent continuous time exploration process. A natural way to study large deviation properties of the configuration model is through the discrete time sequence $\{A(j), \mathbf{V}(j)\}_{j \in \mathbb{N}_0}$ in EEA which can be viewed as a discrete time “small noise” Markov process. In order to study large deviations for such a sequence, a standard approach is to consider the sequence of $\mathbb{C}([0, L] : \mathbb{R}^{\infty})$ -valued random variables obtained by a linear interpolation of $\{A(j), \mathbf{V}(j)\}_{j \in \mathbb{N}_0}$ over intervals of length $1/n$. As was noted in the [Introduction](#), the “diminishing rates” feature of the transition kernel (2.6) makes the large deviations analysis of this sequence challenging. An alternative approach is to consider a continuous time stochastic process that provides a tractable construction of the configuration model. We briefly recall one such construction that was introduced in [31], Section 4.

2.4.1. A simple continuous time construction. In [31], Section 4, it was observed that the configuration model can be explored using a continuous time process constructed using exponential random variables as follows.

1. Every half-edge e is given an independent exponential life-time (call this a clock). Initially, all half-edges and vertices are taken to be sleeping.

2. Whenever the clock of a half-edge rings this half-edge becomes awake and connects to an existing awake half-edge if such a half-edge exists; otherwise it waits for the next half-edge clock to ring and connects to this half-edge completing a full edge. Both such half-edges are then called dead. If at any point a half-edge of a sleeping vertex awakes, that vertex is then said to be awake.

3. The process continues until all half-edges are dead at which point the exploration ends.

It is observed in [31], Section 4, that the random graph constructed at the end of the exploration is a realization from the desired configuration model.

Although the above continuous time construction gives a simple method to produce a sample from the configuration model, it turns out to be hard to directly use it for the study of large deviation problems of interest here. In view of this we present below a different continuous time process for the exploration of the configuration model that is obtained by a more direct Poissonization of the Markov chain $(A(\cdot), V(\cdot))$ in Section 2.3.

2.4.2. A continuous time construction via Poissonization. Let $N(t)$ be a rate- n Poisson process independent of the processes (A, V) of Section 2.3 and define $(\tilde{A}(t), \tilde{V}(t)) \doteq (A(N(t)), V(N(t)))$. Then (\tilde{A}, \tilde{V}) gives a natural continuous time process associated with the exploration of the configuration model. We now give a distributionally equivalent representation of this process which is more tractable for a large deviation analysis. The construction given below ensures that $\{(nX_0^n(\cdot) + 1, nX_k^n(\cdot)), k \in \mathbb{N}\}$, where X_j^n are processes defined below, has the same distribution as $\{\tilde{A}(\cdot), \tilde{V}_k(\cdot), k \in \mathbb{N}\}$.

We begin with some notation that will be needed to formulate the continuous time model. For a locally compact Polish space \mathbb{S} , let $\mathcal{M}_{FC}(\mathbb{S})$ be the space of all measures ν on $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ such that $\nu(K) < \infty$ for every compact $K \subset \mathbb{S}$. We equip $\mathcal{M}_{FC}(\mathbb{S})$ with the usual vague topology. This topology can be metrized such that $\mathcal{M}_{FC}(\mathbb{S})$ is a Polish space (see [15] for one convenient metric). A Poisson random measure (PRM) N on a locally compact Polish space \mathbb{S} with intensity measure $\nu \in \mathcal{M}_{FC}(\mathbb{S})$ is an $\mathcal{M}_{FC}(\mathbb{S})$ -valued random variable such that for each $A \in \mathcal{B}(\mathbb{S})$ with $\nu(A) < \infty$, $N(A)$ is Poisson distributed with mean $\nu(A)$ and for disjoint $A_1, \dots, A_k \in \mathcal{B}(\mathbb{S})$, $N(A_1), \dots, N(A_k)$ are mutually independent random variables (cf. [29]).

Let (Ω, \mathcal{F}, P) be a complete probability space on which are given i.i.d. PRM $\{N_k(ds dy dz)\}_{k \in \mathbb{N}_0}$ on $\mathbb{R}_+ \times [0, 1] \times \mathbb{R}_+$ with intensity measure $ds \times dy \times dz$. Let

$$\hat{\mathcal{F}}_t \doteq \sigma\{N_k((0, s] \times A \times B), 0 \leq s \leq t, A \in \mathcal{B}([0, 1]), B \in \mathcal{B}(\mathbb{R}_+), k \in \mathbb{N}_0\}, \quad t \geq 0$$

and let $\{\mathcal{F}_t\}$ be the P -completion of this filtration. Fix $T \in (0, \infty)$. Let $\bar{\mathcal{P}}$ be the $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ -predictable σ -field on $\Omega \times [0, T]$. Let $\bar{\mathcal{A}}_+$ be all $(\bar{\mathcal{P}} \otimes \mathcal{B}([0, 1]))/\mathcal{B}(\mathbb{R}_+)$ -measurable maps from $\Omega \times [0, T] \times [0, 1]$ to \mathbb{R}_+ . For $\varphi \in \bar{\mathcal{A}}_+$, define a counting process N_k^φ on $[0, T] \times [0, 1]$ by

$$N_k^\varphi([0, t] \times A) \doteq \int_{[0, t] \times A \times \mathbb{R}_+} 1_{[0, \varphi(s, y)]}(z) N_k(ds dy dz), \quad t \in [0, T], A \in \mathcal{B}([0, 1]), k \in \mathbb{N}_0.$$

We think of N_k^φ as a controlled random measure, where φ is the control process that produces a thinning of the point process N_k in a random but nonanticipative manner to produce a desired intensity. We will write N_k^φ as N_k^θ if $\varphi \equiv \theta$ for some constant $\theta \in \mathbb{R}_+$. Note that N_k^θ is a PRM on $[0, T] \times [0, 1]$ with intensity $\theta ds \times dy$. For $\mathbf{x} = (x_0, x_1, x_2, \dots) \in \mathbb{R} \times \mathbb{R}_+^\infty$, let

$$(2.7) \quad \begin{aligned} r(\mathbf{x}) &\doteq (x_0)^+ + \sum_{k=1}^{\infty} kx_k, & r_0(\mathbf{x}) &\doteq \frac{(x_0)^+}{r(\mathbf{x})} 1_{\{r(\mathbf{x}) \in (0, \infty)\}}, \\ r_k(\mathbf{x}) &\doteq \frac{kx_k}{r(\mathbf{x})} 1_{\{r(\mathbf{x}) \in (0, \infty)\}}, & k &\in \mathbb{N}. \end{aligned}$$

Note that $\sum_{k \in \mathbb{N}_0} r_k(\mathbf{x}) = 1$ whenever $r(\mathbf{x}) \in (0, \infty)$. Recall that \mathbf{e}_k is the k th unit vector in \mathbb{R}^∞ , $k \in \mathbb{N}_0$. Define the state process $X^n(t) = (X_0^n(t), X_1^n(t), X_2^n(t), \dots)$ with values in $\mathbb{R} \times \mathbb{R}_+^\infty$ as the solution to the following SDE:

$$\begin{aligned}
 (2.8) \quad X^n(t) = X^n(0) &+ \frac{1}{n} \int_{[0,t] \times [0,1]} 1_{\{X_0^n(s-) \geq 0\}} [-2\mathbf{e}_0] 1_{[0,r_0(X^n(s-))]}(y) N_0^n(ds dy) \\
 &+ \sum_{k=1}^{\infty} \frac{1}{n} \int_{[0,t] \times [0,1]} 1_{\{X_0^n(s-) \geq 0\}} [(k-2)\mathbf{e}_0 - \mathbf{e}_k] 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy) \\
 &+ \sum_{k=1}^{\infty} \frac{1}{n} \int_{[0,t] \times [0,1]} 1_{\{X_0^n(s-) < 0\}} [k\mathbf{e}_0 - \mathbf{e}_k] 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy),
 \end{aligned}$$

where $X^n(0) \doteq \frac{1}{n}(-1, n_1, n_2, \dots)$. The existence and uniqueness of solutions to this SDE follows from the summability of $r_k(\cdot)$. Indeed, for each $\mathbf{z} \in \mathbb{R} \times \mathbb{R}_+^\infty$ and $u \in [0, T]$, the process

$$Z^n(u, \mathbf{z}, t) \doteq \frac{1}{n} \int_{(u,t] \times [0,1]} N_0^n(ds dy) + \sum_{k=1}^{\infty} \frac{1}{n} \int_{(u,t] \times [0,1]} 1_{[0,r_k(\mathbf{z})]}(y) N_k^n(ds dy), \quad u < t \leq T$$

satisfies $Z^n(u, \mathbf{z}, T) < \infty$ since $\sum_{k \in \mathbb{N}_0} r_k(\mathbf{z}) \leq 1$. Together with the mutual independence of the PRM $\{N_k(ds dy dz)\}_{k \in \mathbb{N}_0}$ this says that the jump instants of the point process $\{Z^n(u, \mathbf{z}, t)\}_{u < t \leq T}$ can be enumerated as

$$u < \tau_1^n(\mathbf{z}) < \dots < \tau_{k_n}^n(\mathbf{z}) < T,$$

where $k_n = nZ^n(u, \mathbf{z}, T)$. Thus having constructed the solution of (2.8) on $[0, u]$, the solution can be extended to $[0, \tau_1^n(\mathbf{z})]$, where $\mathbf{z} = X^n(u)$, and the unique solution of (2.8) is now obtained by a standard recursive construction from one jump instant to the next. The solution can be written in an explicit form in terms of the atoms of the PRM $\{N_k^n\}$ which also shows that the solution is a measurable function of the driving PRM. It is not difficult to see that $\frac{1}{n}(A(j) - 1, V_1(j), V_2(j), \dots)$ in the discrete time EEA can be viewed as the embedded Markov chain associated with X^n . Namely, denoting the jump instants of the process X^n as $\{\sigma_j^n\}$, the collection $\{(nX_0^n(\sigma_j^n) + 1, nX_k^n(\sigma_j^n)), k, j \in \mathbb{N}\}$ has the same distribution as $\{A(j), V_k(j), k, j \in \mathbb{N}\}$. In particular, for $k \in \mathbb{N}$, $nX_k^n(\sigma_j^n)$ can be interpreted as the number of sleeping vertices with degree k at the j th step of the exploration in the discrete EEA and in view of Remark 2.14, the excursions of X_0^n away from $-1/n$ track the components in the configuration model. In defining the state process, one could replace $X_0^n(0)$ with the asymptotically equivalent process $X_0^n(0) + 1/n$ which starts from 0 and is more directly comparable with the sequence $A(j)/n$. However some of the expressions are simplified (see, e.g., the formulas for rates in (2.7) and the transition probabilities in (2.6)) when describing the state in terms of $X_0^n(0)$ instead of $X_0^n(0) + 1/n$. We now rewrite the evolution of X^n as follows:

$$\begin{aligned}
 X^n(t) = X^n(0) &+ \mathbf{e}_0 \sum_{k=0}^{\infty} \frac{(k-2)}{n} \int_{[0,t] \times [0,1]} 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy) \\
 &- \sum_{k=1}^{\infty} \mathbf{e}_k \frac{1}{n} \int_{[0,t] \times [0,1]} 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy) \\
 &+ \mathbf{e}_0 \sum_{k=0}^{\infty} \frac{2}{n} \int_{[0,t] \times [0,1]} 1_{\{X_0^n(s-) < 0\}} 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy).
 \end{aligned}$$

Here the first two integrands do not depend on the sign of X_0^n and are interpreted as the main contribution to the evolution. The last sum is a “reflection” term in the e_0 direction and makes a contribution of $\frac{2}{n}e_0$ only when $X_0^n(s-) < 0$. For $t \geq 0$ define

$$(2.9) \quad Y^n(t) \doteq X_0^n(0) + \sum_{k=0}^{\infty} \frac{k-2}{n} \int_{[0,t] \times [0,1]} 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy),$$

$$(2.10) \quad \eta^n(t) \doteq \sum_{k=0}^{\infty} \frac{2}{n} \int_{[0,t] \times [0,1]} 1_{\{X_0^n(s-) < 0\}} 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy).$$

Using these we can write

$$(2.11) \quad X_0^n(t) = Y^n(t) + \eta^n(t),$$

$$(2.12) \quad X_k^n(t) = X_k^n(0) - \frac{1}{n} \int_{[0,t] \times [0,1]} 1_{[0,r_k(X^n(s-))]}(y) N_k^n(ds dy), \quad k \in \mathbb{N}.$$

Here η^n is viewed as the regulator function which ensures that $X_0^n(t) \geq -\frac{1}{n}$. Note that for $k \in \mathbb{N}$, $X_k^n(t)$ is nonincreasing and nonnegative. Also, from (2.8) we see that $r(X^n(t))$ is nonincreasing.

2.5. Rate function. The main result of this work gives a large deviation principle for $\{(X^n, Y^n)\}_{n \in \mathbb{N}}$ in the path space $\mathcal{D}_\infty \times \mathcal{D}$. In this section we define the associated rate function I_T , where the subscript T makes explicit the fact that the processes $\{(X^n, Y^n)\}_{n \in \mathbb{N}}$ are considered on the time horizon $[0, T]$. Including the process Y^n in the LDP is convenient for obtaining large deviation results, for the degree distribution in giant components, of the form given in Section 10.

Recall the probability distribution $\mathbf{p} \doteq \{p_k\}_{k \in \mathbb{N}}$ introduced in Assumption 2.1. In order to describe the rate function it will be convenient to introduce the Skorohod map. The use of the Skorohod reflection mechanism to describe exploration processes for random graphs goes back to the work of Aldous [1]. In the context of large deviation problems for Erdős–Rényi random graph models it has also been used in [42]. Let $\Gamma: \mathcal{C} \rightarrow \mathcal{C}$ denote the one-dimensional Skorokhod map defined by

$$\Gamma(\psi)(t) \doteq \psi(t) - \inf_{0 \leq s \leq t} \psi(s) \wedge 0, \quad t \in [0, T], \psi \in \mathcal{C}.$$

Let \mathcal{C}_T be the subset of $\mathcal{C}_\infty \times \mathcal{C}$, consisting of those functions (ζ, ψ) such that:

- (a) $\psi(0) = 0$, and ψ is absolutely continuous on $[0, T]$.
- (b) $\zeta_0(t) = \Gamma(\psi)(t)$ for $t \in [0, T]$.
- (c) For each $k \in \mathbb{N}$, $\zeta_k(0) = p_k$, ζ_k is nonincreasing and absolutely continuous and $\zeta_k(t) \geq 0$ for $t \in [0, T]$.

For $(\zeta, \psi) \in (\mathcal{D}_\infty \times \mathcal{D}) \setminus \mathcal{C}_T$, define $I_T(\zeta, \psi) \doteq \infty$. For $(\zeta, \psi) \in \mathcal{C}_T$, define

$$(2.13) \quad I_T(\zeta, \psi) \doteq \inf_{\varphi \in \mathcal{S}_T(\zeta, \psi)} \left\{ \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k(s, y)) ds dy \right\}.$$

Here for $x \geq 0$,

$$(2.14) \quad \ell(x) \doteq x \log x - x + 1,$$

and $\mathcal{S}_T(\zeta, \psi)$ is the set of all sequences of functions $\varphi = (\varphi_k)_{k \in \mathbb{N}_0}, \varphi_k : [0, T] \times [0, 1] \rightarrow \mathbb{R}_+$, such that

$$(2.15) \quad \psi(t) = \sum_{k=0}^{\infty} (k-2) \int_{[0,t] \times [0,1]} 1_{[0,r_k(\zeta(s))]}(y) \varphi_k(s, y) ds dy, \quad t \in [0, T],$$

$$(2.16) \quad \zeta_k(t) = p_k - \int_{[0,t] \times [0,1]} 1_{[0,r_k(\zeta(s))]}(y) \varphi_k(s, y) ds dy, \quad k \in \mathbb{N}, t \in [0, T].$$

REMARK 2.15. Suppose $(\zeta, \psi) \in \mathcal{C}_T$ satisfies (2.15) and (2.16) for some $\varphi \in \mathcal{S}_T(\zeta, \psi)$.

(a) From Assumptions 2.1 and 2.2 it follows that the following uniform integrability holds: As $K \rightarrow \infty$:

$$\sup_{0 \leq t \leq T} \sum_{k=K}^{\infty} k \zeta_k(t) \leq \sum_{k=K}^{\infty} k \sup_{0 \leq t \leq T} \zeta_k(t) = \sum_{k=K}^{\infty} k p_k \rightarrow 0.$$

This in particular says that $r(\zeta(\cdot)) \in \mathcal{C}$, where $r(\cdot)$ is defined in (2.7).

(b) For any $k \in \mathbb{N}$, whenever $\zeta_k(t_k) = 0$ for some $t_k \in [0, T]$, we must have $\zeta_k(t) = 0$ for all $t \in [t_k, T]$. This follows since ζ_k is nonincreasing and nonnegative for every k .

(c) Whenever $r(\zeta(t^*)) = 0$ for some $t^* \in [0, T]$, we must have from part (b) that $\zeta_k(t) = 0$ for all $t \in [t^*, T]$ and $k \in \mathbb{N}$. This, together with (2.15), implies that $\psi(\cdot)$ is nonincreasing on the interval $[t^*, T]$. Hence by property (b) of \mathcal{C}_T , $\zeta_0(t)$ is nonincreasing and nonnegative for $t \in [t^*, T]$. Since $\zeta_0(t^*) = 0$, we must then have $\zeta_0(t) = 0$ for $t \in [t^*, T]$, which means that $\zeta(t) = \mathbf{0}$ for $t \in [t^*, T]$. Thus whenever such a t^* exists, $\zeta(t) = \mathbf{0}$ after the time instant

$$(2.17) \quad \tau_{\zeta} \doteq \inf\{t \in [0, T] : r(\zeta(t)) = 0\} \wedge T.$$

2.6. *LDP and LLN for the exploration process.* The following LDP is one of our main results and is key to the proof of Theorem 2.6.

THEOREM 2.16. *The function I_T in (2.13) is a rate function on $\mathcal{D}_{\infty} \times \mathcal{D}$ and the sequence $\{(X^n, Y^n)\}_{n \in \mathbb{N}}$ satisfies a large deviation principle in $\mathcal{D}_{\infty} \times \mathcal{D}$ with rate function I_T .*

Outline of the proof. Due to the equivalence between a large deviation principle and a Laplace principle, it suffices to show the following three statements (cf. [25], Section 1.2, or [13], Section 1.2).

(1) Laplace principle upper bound: For all $h \in \mathbb{C}_b(\mathcal{D}_{\infty} \times \mathcal{D})$,

$$(2.18) \quad \limsup_{n \rightarrow \infty} \frac{1}{n} \log E e^{-nh(X^n, Y^n)} \leq - \inf_{(\zeta, \psi) \in \mathcal{C}_{\infty} \times \mathcal{C}} \{I_T(\zeta, \psi) + h(\zeta, \psi)\}.$$

(2) Laplace principle lower bound: For all $h \in \mathbb{C}_b(\mathcal{D}_{\infty} \times \mathcal{D})$,

$$(2.19) \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log E e^{-nh(X^n, Y^n)} \geq - \inf_{(\zeta, \psi) \in \mathcal{C}_{\infty} \times \mathcal{C}} \{I_T(\zeta, \psi) + h(\zeta, \psi)\}.$$

(3) I_T is a rate function on $\mathcal{D}_{\infty} \times \mathcal{D}$: For each $M \in [0, \infty)$, $\{(\zeta, \psi) \in \mathcal{D}_{\infty} \times \mathcal{D} : I_T(\zeta, \psi) \leq M\}$ is a compact subset of $\mathcal{D}_{\infty} \times \mathcal{D}$.

Statements (1), (2) and (3) will be shown in Sections 4, 5 and 6, respectively.

REMARK 2.17. As noted above, the LDP in Theorem 2.16 is a key to the proof of Theorem 2.6. In the next subsection we will show how this LDP can be used to easily give a LLN result. The LDP can be used to establish other asymptotic results as well. We give one

such example without proof below. Denote by C^n the number of components in $G([n], \mathbf{d}(n))$. Then η^n defined in (2.10) can be used to represent $\frac{C^n}{n}$. Such an observation in the context of Erdős-Rényi random graphs was first made in [1] (see also [42]). Note that whenever the EEA starts to explore a new component, X_0^n will jump from $-\frac{1}{n}$ and as a result, η^n will increase by $\frac{2}{n}$. Therefore,

$$\frac{C^n}{n} = \sup_{t>0} \frac{\eta^n(t)}{2} = \lim_{T \rightarrow \infty} \frac{\eta^n(T)}{2}.$$

Observe from (2.11) that $\eta^n = X_0^n - Y^n$, and that for large deviation asymptotics one can assume that the EEA terminates by time $N \doteq 1 + \lfloor \sup_n \frac{1}{2} \sum_{k=1}^{\infty} k \frac{n_k^{(n)}}{n} \rfloor < \infty$ (see Lemma 8.1 and its proof for precise details). Using this fact, Theorem 2.16, and the contraction principle one can establish that $\frac{C^n}{n}$ satisfies a large deviation principle in \mathbb{R}_+ with rate function \hat{I} defined by

$$\hat{I}(x) = \lim_{T \rightarrow \infty} \inf_{(\xi, \psi) \in \mathcal{C}_T : \xi_0(T) - \psi(T) = 2x} I_T(\xi, \psi).$$

The rate function $\hat{I}(x)$ has the following alternative representation.

$$\begin{aligned} \hat{I}(x) = & \inf_{(\xi, \psi) \in \mathcal{C}_N : \xi_0(N) - \psi(N) = 2x, \xi(N) = 0} \int_0^N \left[r_0(\xi(t)) \ell \left(-\frac{\psi'(t) + \sum_{k=1}^{\infty} (k-2) \xi'_k(t)}{2r_0(\xi(t))} \right) \right. \\ & \left. + \sum_{k=1}^{\infty} r_k(\xi(t)) \ell \left(-\frac{\xi'_k(t)}{r_k(\xi(t))} \right) \right] dt. \end{aligned}$$

2.6.1. Law of large number limits. The LDP in Theorem 2.16 can be used to identify the LLN limit (ξ, ψ) of the exploration process (X^n, Y^n) , which corresponds to the *unique* pair satisfying $I_T(\xi, \psi) = 0$. In particular we recover well-known results for the asymptotics of the largest component in the configuration model [31, 37]. We assume the following strengthened version of Assumption 2.2.

ASSUMPTION 2.18. $\sup_{n \in \mathbb{N}} \sum_{k=1}^{\infty} \frac{n_k}{n} k^2 < \infty$.

REMARK 2.19. Under our standing assumptions, namely Assumption 2.1 and 2.2, one can show by following the arguments in Section 3 that $\{(X^n, Y^n)\}_{n \in \mathbb{N}}$ is tight and any weak limit point (ξ, ψ) of this sequence is in \mathcal{C}_T and satisfies (2.15) and (2.16) with $\varphi_k = 1$ for $k \in \mathbb{N}_0$. However it seems hard to argue the uniqueness of this limiting system of equations without additional conditions. Instead we show that if Assumption 2.2 is replaced with the stronger condition in Assumption 2.18 then there is an explicit trajectory (ξ, ψ) for which the rate function vanishes and in fact it is the unique such trajectory. This is the content of Theorem 2.20 and Proposition 2.21. From these results the LLN follows immediately. Whether the LLN holds under the weaker Assumption 2.2 is an open problem.

Recall $\mu \doteq \sum_{k=1}^{\infty} k p_k$ and note that $\mu < \infty$. Define, for $z \in [0, 1]$,

$$G_0(z) \doteq \sum_{k=1}^{\infty} p_k z^k \quad \text{and} \quad G_1(z) \doteq \sum_{k=1}^{\infty} \frac{k p_k}{\mu} z^{k-1}.$$

Define $F_s(t) \doteq G_0(s) - G_0(st)$ for $s \in (0, 1]$ and $t \in [0, 1]$. Then $F_s : [0, 1] \rightarrow [0, G_0(s)]$ is strictly decreasing and continuous. Let $F_s^{-1}(\cdot)$ denote the inverse of F_s . Define

$$f_s(t) \doteq \begin{cases} F_s^{-1}(t) & \text{when } 0 \leq t \leq G_0(s), \\ 0 & \text{when } t > G_0(s). \end{cases}$$

Then $f_s(t)$ is strictly decreasing until it hits zero. Note that in particular, $f_1(t) = F_1^{-1}(t) \times 1_{[0,1]}(t)$. Define $f_0(t) \doteq 0$ for $t \geq 0$.

Fix $T \geq \frac{\mu}{2}$. The following theorem together with Proposition 2.21 characterizes the unique $(\zeta, \psi) \in \mathcal{C}_T$ that minimizes the rate function $I_T(\zeta, \psi)$. Letting

$$\nu \doteq \frac{\sum_{k=1}^{\infty} k(k-1)p_k}{\sum_{k=1}^{\infty} kp_k},$$

part 1 of the theorem considers the subcritical and critical cases $\nu \leq 1$, where the size of the largest component is $o(n)$, while part 2 considers the supercritical case $\nu > 1$, where the size of the largest component is $O(n)$. Proofs of Theorem 2.20 and Proposition 2.21 are provided in Section 10.

THEOREM 2.20. *Suppose that Assumptions 2.1 and 2.18 hold.*

(1) *Suppose $\sum_{k=1}^{\infty} k(k-2)p_k \leq 0$. Define $\zeta(t) = (\zeta_k(t))_{k \in \mathbb{N}_0}$ and $\psi(t)$ by*

$$\zeta_0(t) \doteq 0, \zeta_k(t) \doteq p_k(f_1(t))^k, \quad k \in \mathbb{N},$$

$$\psi(t) \doteq -2 \int_0^t r_0(\zeta(s)) ds + \sum_{k=1}^{\infty} (k-2)(p_k - \zeta_k(t)).$$

Then $(\zeta, \psi) \in \mathcal{C}_T$ and $I_T(\zeta, \psi) = 0$.

(2) *Suppose $\sum_{k=1}^{\infty} k(k-2)p_k > 0$. If $p_1 > 0$, then there exists a unique $\rho \in (0, 1)$ such that $G_1(\rho) = \rho$. If $p_1 = 0$, $G_1(\rho) = \rho$ with $\rho \doteq 0$. Define $\tau = \frac{\mu}{2}(1 - \rho^2) > 0$ and define $\zeta(t) = (\zeta_k(t))_{k \in \mathbb{N}_0}$ and $\psi(t)$ by*

$$\zeta_0(t) \doteq [\mu - 2t - \mu\sqrt{1-2t/\mu}G_1(\sqrt{1-2t/\mu})]1_{[0,\tau]}(t),$$

$$\zeta_k(t) \doteq \begin{cases} p_k(1-2t/\mu)^{k/2} & \text{when } 0 \leq t \leq \tau, \\ p_k\rho^k(f_\rho(t-\tau))^k & \text{when } t > \tau, \end{cases} \quad k \in \mathbb{N},$$

$$\psi(t) \doteq -2 \int_0^t r_0(\zeta(s)) ds + \sum_{k=1}^{\infty} (k-2)(p_k - \zeta_k(t)).$$

Then $(\zeta, \psi) \in \mathcal{C}_T$ and $I_T(\zeta, \psi) = 0$.

The following proposition says that there is a unique (ζ, ψ) satisfying $I_T(\zeta, \psi) = 0$, so that this pair is the law of large numbers limit.

PROPOSITION 2.21. *Suppose Assumptions 2.1 and 2.18 hold. Then the pair (ζ, ψ) defined in Theorem 2.20 is the unique element of $\mathcal{D}_\infty \times \mathcal{D}$ such that $I_T(\zeta, \psi) = 0$.*

3. Representation and weak convergence of controlled processes. We will use the following useful representation formula proved in [15]. For the second equality in the theorem see the proof of Theorem 2.4 in [11]. The representation in the cited papers is given in terms of a single Poisson random measure with points in a locally compact Polish space. However for the current work it is convenient to formulate the representation in terms of a countable sequence of independent Poisson random measures on $[0, T] \times [0, 1]$. This representation is immediate from the results in [15] and [11] by viewing the countable sequence of Poisson random measures with points in $[0, T] \times [0, 1]$ and intensity the Lebesgue measure λ_T on $[0, T] \times [0, 1]$ as a single PRM with points in the augmented space $[0, T] \times [0, 1] \times \mathbb{N}_0$ with intensity $\lambda_T \otimes \varrho$, where ϱ is the counting measure on \mathbb{N} . Recall that \mathcal{A}_+ denotes the class of

$(\bar{\mathcal{P}} \times \mathcal{B}([0, 1]))/\mathcal{B}(\mathbb{R}_+)$ -measurable maps from $\Omega \times [0, T] \times [0, 1]$ to \mathbb{R}_+ . For each $m \in \mathbb{N}$ let

$$\bar{\mathcal{A}}_{b,m} \doteq \{(\varphi_k)_{k \in \mathbb{N}_0} : \varphi_k \in \bar{\mathcal{A}}_+ \text{ for each } k \in \mathbb{N}_0 \text{ such that for all } (\omega, t, y) \in \Omega \times [0, T] \times [0, 1], \\ 1/m \leq \varphi_k(\omega, t, y) \leq m \text{ for } k \leq m \text{ and } \varphi_k(\omega, t, y) = 1 \text{ for } k > m\}$$

and let $\bar{\mathcal{A}}_b \doteq \bigcup_{m=1}^{\infty} \bar{\mathcal{A}}_{b,m}$. Recall the function ℓ defined in (2.14).

THEOREM 3.1. *Let $F \in \mathbb{M}_b([\mathcal{M}_{FC}([0, T] \times [0, 1])]^\infty)$. Then, for $\theta > 0$,*

$$\begin{aligned} & -\log E e^{-F((N_k^\theta)_{k \in \mathbb{N}_0})} \\ &= \inf_{\varphi_k \in \bar{\mathcal{A}}_+, k \in \mathbb{N}_0} E \left[\theta \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy + F((N_k^{\theta \varphi_k})_{k \in \mathbb{N}_0}) \right] \\ &= \inf_{\varphi = (\varphi_k)_{k \in \mathbb{N}_0} \in \bar{\mathcal{A}}_b} E \left[\theta \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy + F((N_k^{\theta \varphi_k})_{k \in \mathbb{N}_0}) \right]. \end{aligned}$$

Fix $h \in \mathbb{C}_b(\mathcal{D}_\infty \times \mathcal{D})$. Since (X^n, Y^n) can be written as $\Psi((N_k^n)_{k \in \mathbb{N}_0})$ for some measurable function Ψ from $[\mathcal{M}_{FC}([0, T] \times [0, 1])]^\infty$ to $\mathcal{D}_\infty \times \mathcal{D}$, we have from the second equality in Theorem 3.1 that, with $(\theta, F) = (n, nh \circ \Psi)$,

$$\begin{aligned} & -\frac{1}{n} \log E e^{-nh(X^n, Y^n)} \\ (3.1) \quad &= \inf_{\varphi^n = (\varphi_k^n)_{k \in \mathbb{N}_0} \in \bar{\mathcal{A}}_b} E \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy + h(\bar{X}^n, \bar{Y}^n) \right\}. \end{aligned}$$

Here $(\bar{X}^n, \bar{Y}^n) = \Psi((N_k^{n\varphi_k^n})_{k \in \mathbb{N}_0})$, which solves the controlled analogue of (2.9)–(2.12), namely $\bar{X}^n(0) \doteq \frac{1}{n}(-1, n_1, n_2, \dots)$, and for $t \in [0, T]$,

$$(3.2) \quad \bar{Y}^n(t) = \bar{X}_0^n(0) + \sum_{k=0}^{\infty} \frac{k-2}{n} \int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}^n(s-))]}(y) N_k^{n\varphi_k^n}(ds dy),$$

$$(3.3) \quad \bar{X}_0^n(t) = \bar{Y}^n(t) + \frac{2}{n} \sum_{k=0}^{\infty} \int_{[0, t] \times [0, 1]} 1_{\{\bar{X}_0^n(s-) < 0\}} 1_{[0, r_k(\bar{X}^n(s-))]}(y) N_k^{n\varphi_k^n}(ds dy),$$

$$(3.4) \quad \bar{X}_k^n(t) = \bar{X}_k^n(0) - \frac{1}{n} \int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}^n(s-))]}(y) N_k^{n\varphi_k^n}(ds dy), \quad k \in \mathbb{N}.$$

There is a bar in the notation \bar{X}^n, \bar{Y}^n (and \bar{v}^n defined in (3.12) below) to indicate that these are “controlled” processes, given in terms of the control sequence $\varphi^n \doteq (\varphi_k^n)_{k \in \mathbb{N}_0}$. We will occasionally suppress the dependence on φ^n in the notation and will make this dependence explicit if there are multiple controls (e.g., as in Section 4).

In the proof of both the upper and lower bound we will show it is sufficient to consider a sequence $\{\varphi_k^n \in \bar{\mathcal{A}}_+, k \in \mathbb{N}_0\}$ that satisfies the following uniform bound for some $M_0 < \infty$:

$$(3.5) \quad \sup_{n \in \mathbb{N}} \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy \leq M_0, \quad \text{a.s. } P.$$

In the rest of this section we study tightness and convergence properties of controlled processes (\bar{X}^n, \bar{Y}^n) that are driven by controls $\{\varphi_k^n\}$ that satisfy the above a.s. bound.

From (3.2)–(3.4) we can rewrite

$$(3.6) \quad \bar{Y}^n(t) = \bar{X}_0^n(0) + \sum_{k=0}^{\infty} (k-2) \bar{B}_k^n(t),$$

$$(3.7) \quad \bar{X}_0^n(t) = \bar{Y}^n(t) + \bar{\eta}^n(t),$$

$$(3.8) \quad \bar{X}_k^n(t) = \bar{X}_k^n(0) - \bar{B}_k^n(t), \quad k \in \mathbb{N},$$

where

$$(3.9) \quad \bar{B}_k^n(t) \doteq \frac{1}{n} \int_{[0,t] \times [0,1]} 1_{[0,r_k(\bar{X}^n(s-))]}(y) N_k^{n\varphi_k^n}(ds dy), \quad k \in \mathbb{N}_0,$$

$$(3.10) \quad \begin{aligned} \bar{\eta}^n(t) &\doteq \sum_{k=0}^{\infty} \frac{2}{n} \int_{[0,t] \times [0,1]} 1_{\{\bar{X}_0^n(s-) < 0\}} 1_{[0,r_k(\bar{X}^n(s-))]}(y) N_k^{n\varphi_k^n}(ds dy) \\ &= \sum_{k=1}^{\infty} \frac{2}{n} \int_{[0,t] \times [0,1]} 1_{\{\bar{X}_0^n(s-) < 0\}} 1_{[0,r_k(\bar{X}^n(s-))]}(y) N_k^{n\varphi_k^n}(ds dy). \end{aligned}$$

Here the last line follows on observing that $1_{\{\bar{X}_0^n(s-) < 0\}} 1_{[0,r_0(\bar{X}^n(s-))]}(y) \equiv 0$.

Since $m_1 \doteq \sup_{n \in \mathbb{N}} \sum_{k=1}^{\infty} k \frac{n_k}{n} < \infty$ by Assumption 2.2, using (2.7) we have $-\frac{1}{n} \leq \bar{X}_0^n(t) \leq m_1$, $0 \leq r(\bar{X}^n(t)) \leq m_1$ and $0 \leq \bar{X}_k^n(t) \leq \frac{n_k}{n}$ for $t \in [0, T]$. In particular, the nonnegativity of $\bar{X}_k^n(t)$ is an immediate consequence of the evolution equation (3.4) on observing that $r_k(\bar{X}^n(s-)) = 0$ if $\bar{X}_k^n(s-) = 0$ and that the jumps of \bar{X}_k^n are of size $1/n$. Also note that both $r(\bar{X}^n(\cdot))$ and $\bar{X}_k^n(\cdot)$ for $k \in \mathbb{N}$ are nonincreasing.

The following lemma summarizes some elementary properties of ℓ . For part (a) we refer to [14], Lemma 3.1, and part (b) is an easy calculation that is omitted.

LEMMA 3.2.

- (a) For each $\beta > 0$, there exists $\gamma(\beta) \in (0, \infty)$ such that $\gamma(\beta) \rightarrow 0$ as $\beta \rightarrow \infty$ and $x \leq \gamma(\beta)\ell(x)$, for $x \geq \beta > 1$.
- (b) For $x \geq 0$, $x \leq \ell(x) + 2$.

The next lemma gives some uniform integrability properties for the control sequence φ^n .

LEMMA 3.3. For $K \in \mathbb{N}$ define

$$(3.11) \quad \bar{U}_K \doteq \sup_{n \in \mathbb{N}} E \left\{ \sum_{k=K}^{\infty} \int_{[0,T] \times [0,1]} k \varphi_k^n(s, y) 1_{[0,r_k(\bar{X}^n(s))]}(y) ds dy \right\}.$$

Then $\bar{U}_K < \infty$ for each $K \in \mathbb{N}$ and $\lim_{K \rightarrow \infty} \bar{U}_K = 0$.

PROOF. From (3.9) and (3.8) it follows that

$$\bar{U}_K = \sup_{n \in \mathbb{N}} E \left\{ \sum_{k=K}^{\infty} k \bar{B}_k^n(T) \right\} = \sup_{n \in \mathbb{N}} E \left\{ \sum_{k=K}^{\infty} k [\bar{X}_k^n(0) - \bar{X}_k^n(T)] \right\} \leq \sup_{n \in \mathbb{N}} \sum_{k=K}^{\infty} \frac{k n_k}{n}.$$

Recalling $\varepsilon_p \in (0, \infty)$ introduced in Assumption 2.2, we have

$$\sup_{n \in \mathbb{N}} \sum_{k=K}^{\infty} \frac{k n_k}{n} \leq K^{-\varepsilon_p} \sup_{n \in \mathbb{N}} \sum_{k=1}^{\infty} \frac{n_k}{n} k^{1+\varepsilon_p} \rightarrow 0$$

as $K \rightarrow \infty$. The result follows. \square

The following lemma proves some key tightness properties. Write $\bar{\mathbf{B}}^n \doteq \{\bar{B}_k^n\}_{k \in \mathbb{N}_0}$. Define $\bar{\mathbf{v}}^n \doteq \{\bar{v}_k^n\}_{k \in \mathbb{N}_0}$, where for $k \in \mathbb{N}_0$,

$$(3.12) \quad \bar{v}_k^n([0, t] \times A) \doteq \int_{[0, t] \times A} \varphi_k^n(s, y) ds dy, \quad t \in [0, T], A \in \mathcal{B}([0, 1]).$$

LEMMA 3.4. *Suppose that the bound in (3.5) is satisfied. Then the sequence of random variables $\{(\bar{\mathbf{v}}^n, \bar{X}^n, \bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n)\}$ is tight in $[\mathcal{M}([0, T] \times [0, 1])]^\infty \times \mathcal{D}_\infty \times \mathcal{D} \times \mathcal{D}_\infty \times \mathcal{D}$.*

PROOF. We will argue the tightness of $\{\bar{\mathbf{v}}^n\}$ in $[\mathcal{M}([0, T] \times [0, 1])]^\infty$ and the \mathcal{C} -tightness of $\{\bar{X}^n\}$, $\{\bar{Y}^n\}$, $\{\bar{\mathbf{B}}^n\}$, and $\{\bar{\eta}^n\}$ in \mathcal{D}_∞ , \mathcal{D} , \mathcal{D}_∞ , and \mathcal{D} respectively. This will complete the proof.

Consider first $\{\bar{\mathbf{v}}^n\}$. Note that $[0, T] \times [0, 1]$ is a compact metric space. Also from Lemma 3.2(b) and (3.5) we have a.s. for each $k \in \mathbb{N}_0$,

$$\begin{aligned} \bar{v}_k^n([0, T] \times [0, 1]) &= \int_{[0, T] \times [0, 1]} \varphi_k^n(s, y) ds dy \\ &\leq \int_{[0, T] \times [0, 1]} (\ell(\varphi_k^n(s, y)) + 2) ds dy \leq M_0 + 2T. \end{aligned}$$

Hence $\{\bar{v}_k^n\}$ is tight in $\mathcal{M}([0, T] \times [0, 1])$.

Next, since $\bar{X}_k^n(0) \in [0, 1]$ for $k \in \mathbb{N}$ a.s., we see from (3.7) and (3.8) that \mathcal{C} -tightness of $\{\bar{X}^n\}$ in \mathcal{D}_∞ follows once we show \mathcal{C} -tightness of $\{\bar{Y}^n\}$, $\{\bar{\mathbf{B}}^n\}$ and $\{\bar{\eta}^n\}$.

We now show that $\{(\bar{Y}^n(t), \bar{\mathbf{B}}^n(t), \bar{\eta}^n(t))\}$ is tight for each t . From (3.6), (3.9) and (3.10),

$$\begin{aligned} &E \left[|\bar{Y}^n(t)| + \sum_{k=0}^{\infty} |\bar{B}_k^n(t)| + |\bar{\eta}^n(t)| \right] \\ &\leq \frac{1}{n} + \sum_{k=0}^{\infty} [|k-2|+1] E |\bar{B}_k^n(t)| + E |\bar{\eta}^n(t)| \\ &\leq \frac{1}{n} + E \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} [|k-2|+1 + 2 \cdot 1_{\{k \geq 1\}}] \varphi_k^n(s, y) 1_{[0, r_k(\bar{X}^n(s))]}(y) ds dy \\ &\leq \frac{1}{n} + 3E \int_{[0, T] \times [0, 1]} \varphi_0^n(s, y) ds dy + 4\bar{U}_1, \end{aligned}$$

where the last line uses the definition of \bar{U}_1 in (3.11). From Lemma 3.2(b) and (3.5), we have

$$E \int_{[0, T] \times [0, 1]} \varphi_0^n(s, y) ds dy \leq E \int_{[0, T] \times [0, 1]} [\ell(\varphi_0^n(s, y)) + 2] ds dy \leq M_0 + 2T.$$

Therefore, $\sup_{n \in \mathbb{N}} E[|\bar{Y}^n(t)| + \sum_{k=0}^{\infty} |\bar{B}_k^n(t)| + |\bar{\eta}^n(t)|] < \infty$ and we have tightness of $\{(\bar{Y}^n(t), \bar{\mathbf{B}}^n(t), \bar{\eta}^n(t))\}$ in $\mathbb{R} \times \mathbb{R}^\infty \times \mathbb{R}$ for each $t \in [0, T]$.

We now consider fluctuations of $(\bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n)$. Recall the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$. For $\delta \in [0, T]$, let \mathcal{T}^δ be the collection of all $[0, T-\delta]$ -valued stopping times τ . Note that for $\tau \in \mathcal{T}^\delta$,

$$E |\bar{Y}^n(\tau + \delta) - \bar{Y}^n(\tau)| \leq E \left[\sum_{k=0}^{\infty} (k+2) |\bar{B}_k^n(\tau + \delta) - \bar{B}_k^n(\tau)| \right].$$

Thus in order to argue tightness of $\{(\bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n)\}$, by the Aldous–Kurtz tightness criterion (cf. [33], Theorem 2.7) it suffices to show that

$$(3.13) \quad \limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{T}^\delta} E \left[\sum_{k=0}^{\infty} (k+2) |\bar{B}_k^n(\tau + \delta) - \bar{B}_k^n(\tau)| + |\bar{\eta}^n(\tau + \delta) - \bar{\eta}^n(\tau)| \right] = 0.$$

From (3.9) and (3.10) it follows that for every $K \in \mathbb{N}$ and $M \in (0, \infty)$,

$$\begin{aligned} & E \left[\sum_{k=0}^{\infty} (k+2) |\bar{B}_k^n(\tau + \delta) - \bar{B}_k^n(\tau)| + |\bar{\eta}^n(\tau + \delta) - \bar{\eta}^n(\tau)| \right] \\ & \leq E \sum_{k=0}^{\infty} \int_{(\tau, \tau + \delta] \times [0, 1]} (k+4) \varphi_k^n(s, y) 1_{[0, r_k(\bar{X}^n(s))]}(y) ds dy \\ & \leq E \sum_{k=0}^{K-1} \left[\int_{(\tau, \tau + \delta] \times [0, 1]} (k+4) \varphi_k^n(s, y) 1_{\{\varphi_k^n(s, y) > M\}} ds dy \right. \\ & \quad \left. + \int_{(\tau, \tau + \delta] \times [0, 1]} (k+4) \varphi_k^n(s, y) 1_{\{\varphi_k^n(s, y) \leq M\}} ds dy \right] + 5\bar{U}_K. \end{aligned}$$

Using Lemma 3.2(a) and (3.5), we can bound the last display by

$$\begin{aligned} & E \sum_{k=0}^{K-1} \int_{(\tau, \tau + \delta] \times [0, 1]} (K+3) \gamma(M) \ell(\varphi_k^n(s, y)) ds dy + K(K+3)M\delta + 5\bar{U}_K \\ & \leq (K+3)\gamma(M)M_0 + K(K+3)M\delta + 5\bar{U}_K. \end{aligned}$$

Therefore,

$$\begin{aligned} & \limsup_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\tau \in \mathcal{T}^\delta} E \left[\sum_{k=0}^{\infty} (k+2) |\bar{B}_k^n(\tau + \delta) - \bar{B}_k^n(\tau)| + |\bar{\eta}^n(\tau + \delta) - \bar{\eta}^n(\tau)| \right] \\ & \leq (K+3)\gamma(M)M_0 + 5\bar{U}_K. \end{aligned}$$

Taking $M \rightarrow \infty$ and then $K \rightarrow \infty$, we have from Lemmas 3.2(a) and 3.3 that (3.13) holds. Finally \mathcal{C} -tightness is immediate from the following a.s. bounds, Assumption 2.2, and [5], Theorem 13.4: for any $k \in \mathbb{N}_0$, $K \in \mathbb{N}$ and $t \in (0, T]$,

$$\begin{aligned} |\bar{B}_k^n(t) - \bar{B}_k^n(t-)| & \leq \frac{1}{n}, \\ |\bar{\eta}^n(t) - \bar{\eta}^n(t-)| & \leq \frac{2}{n}, \\ |\bar{Y}_k^n(t) - \bar{Y}_k^n(t-)| & \leq \frac{K}{n} + \sum_{j=K+1}^{\infty} \frac{j n_j}{n}. \end{aligned}$$

This completes the proof. \square

Next we will characterize weak limit points of $\{(\bar{v}^n, \bar{X}^n, \bar{Y}^n, \bar{B}^n, \bar{\eta}^n)\}$. For that, we need the following notation. For $k \in \mathbb{N}_0$ define the compensated process

$$\tilde{N}_k^{n\varphi_k^n}(ds dy) \doteq N_k^{n\varphi_k^n}(ds dy) - n\varphi_k^n(s, y) ds dy.$$

Then $\tilde{N}_k^{n\varphi_k^n}([0, t] \times A)$ is an $\{\mathcal{F}_t\}$ -martingale for $A \in \mathcal{B}([0, 1])$ and $k \in \mathbb{N}_0$. Let

$$(3.14) \quad \tilde{B}_k^n(t) = \tilde{B}_k^n(t) + \hat{B}_k^n(t), \quad t \in [0, T], k \in \mathbb{N}_0$$

where

$$\tilde{B}_k^n(t) \doteq \frac{1}{n} \int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}^n(s-))]}(y) \tilde{N}_k^{n\varphi_k^n}(ds dy)$$

is an $\{\mathcal{F}_t\}$ -martingale and

$$\hat{B}_k^n(t) \doteq \int_{[0,t] \times [0,1]} 1_{[0,r_k(\bar{X}^n(s))]}(y) \varphi_k^n(s, y) ds dy.$$

Write $\tilde{\mathbf{B}}^n \doteq (\tilde{B}_k^n)_{k \in \mathbb{N}_0}$ and $\hat{\mathbf{B}}^n \doteq (\hat{B}_k^n)_{k \in \mathbb{N}_0}$. Let λ_t be Lebesgue measure on $[0, t] \times [0, 1]$.

We have the following characterization of the weak limit points. Recall $\mathcal{S}_T(\xi, \psi)$ defined in (2.15) and (2.16).

LEMMA 3.5. *Suppose Assumptions 2.1 and 2.2 hold. Also assume that the bound (3.5) is satisfied and suppose that $(\bar{\mathbf{v}}^n, \bar{X}^n, \bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n)$ converges along a subsequence, in distribution, to $(\bar{\mathbf{v}}, \bar{X}, \bar{Y}, \bar{\mathbf{B}}, \bar{\eta}) \in [\mathcal{M}([0, T] \times [0, 1])]^\infty \times \mathcal{D}_\infty \times \mathcal{D} \times \mathcal{D}_\infty \times \mathcal{D}$ given on some probability space $(\Omega^*, \mathcal{F}^*, P^*)$. Then the following holds P^* -a.s.*

- (a) *For each $k \in \mathbb{N}_0$, $\bar{v}_k \ll \lambda_T$.*
- (b) *$(\bar{X}, \bar{Y}, \bar{\mathbf{B}}, \bar{\eta}) \in \mathcal{C}_\infty \times \mathcal{C} \times \mathcal{C}_\infty \times \mathcal{C}$, and for $t \in [0, T]$*

$$(3.15) \quad \bar{X}_k(t) = p_k - \bar{B}_k(t) \geq 0, \quad k \in \mathbb{N},$$

$$(3.16) \quad \bar{Y}(t) = \sum_{k=0}^{\infty} (k-2) \bar{B}_k(t),$$

$$(3.17) \quad \bar{X}_0(t) = \bar{Y}(t) + \bar{\eta}(t) \geq 0.$$

- (c) *For $k \in \mathbb{N}_0$ let $\varphi_k(s, y) \doteq \frac{d\bar{v}_k}{d\lambda_T}(s, y)$, $(s, y) \in [0, T] \times [0, 1]$. Then for $t \in [0, T]$ and $k \in \mathbb{N}_0$*

$$(3.18) \quad \bar{B}_k(t) = \int_{[0,t] \times [0,1]} 1_{[0,r_k(\bar{X}(s))]}(y) \varphi_k(s, y) ds dy.$$

- (d) *$\bar{X}_0 = \Gamma(\bar{Y})$. In particular, $(\bar{X}, \bar{Y}) \in \mathcal{C}_T$ and $\varphi \in \mathcal{S}_T(\bar{X}, \bar{Y})$.*

PROOF. Assume without loss of generality that $(\bar{\mathbf{v}}^n, \bar{X}^n, \bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n) \Rightarrow (\bar{\mathbf{v}}, \bar{X}, \bar{Y}, \bar{\mathbf{B}}, \bar{\eta})$ along the whole sequence as $n \rightarrow \infty$.

- (a) This is an immediate consequence of the bound in (3.5) and Lemma A.1 of [11].
- (b) The first statement is an immediate consequence of the \mathcal{C} -tightness argued in the proof of Lemma 3.4. Then using (3.8), Assumption 2.1 and the fact that $\bar{X}_k^n(t) \geq 0$ a.s., we have (3.15). Next, note that by Assumption 2.2, as $K \rightarrow \infty$

$$(3.19) \quad \sup_{n \in \mathbb{N}} \sup_{0 \leq t \leq T} \left| \sum_{k=K}^{\infty} (k-2) \bar{B}_k^n(t) \right| \leq \sup_{n \in \mathbb{N}} \sum_{k=K}^{\infty} \frac{kn_k}{n} \leq K^{-\varepsilon_p} \sup_{n \in \mathbb{N}} \sum_{k=K}^{\infty} \frac{n_k}{n} k^{1+\varepsilon_p} \rightarrow 0,$$

where in obtaining the first inequality we have used the fact that due to (3.8) and the nonnegativity of $\bar{X}_k^n(t)$, $\bar{B}_k^n(t) \leq \bar{X}_k^n(0)$. Hence $\sum_{k=0}^{\infty} (k-2) \bar{B}_k^n \Rightarrow \sum_{k=0}^{\infty} (k-2) \bar{B}_k \in \mathcal{C}$. From this and (3.6) we see that (3.16) holds. Next, since $(\bar{Y}^n, \bar{\eta}^n) \Rightarrow (\bar{Y}, \bar{\eta}) \in \mathcal{C}^2$ and $\bar{X}_0^n(t) \geq -\frac{1}{n}$ a.s., we have from (3.7) that (3.17) holds.

- (c) By Doob's inequality, as $n \rightarrow \infty$

$$\begin{aligned} E \sum_{k=0}^{\infty} \sup_{0 \leq t \leq T} |\tilde{B}_k^n(t)|^2 &\leq \frac{4}{n} E \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \varphi_k^n(s, y) 1_{[0,r_k(\bar{X}^n(s))]}(y) ds dy \\ &\leq \frac{4}{n} E \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} [\ell(\varphi_k^n(s, y)) + 2] 1_{[0,r_k(\bar{X}^n(s))]}(y) ds dy \\ &\leq \frac{4}{n} (M_0 + 2T) \rightarrow 0, \end{aligned}$$

where the second inequality follows from Lemma 3.2(b) and the third inequality follows from (3.5). Therefore, as $n \rightarrow \infty$

$$(3.20) \quad \tilde{\mathbf{B}}^n \Rightarrow \mathbf{0}.$$

By appealing to the Skorokhod representation theorem, we can assume without loss of generality that $(\bar{v}^n, \bar{X}^n, \bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n, \tilde{\mathbf{B}}^n) \rightarrow (\bar{v}, \bar{X}, \bar{Y}, \bar{\mathbf{B}}, \bar{\eta}, \mathbf{0})$ a.s. on $(\Omega^*, \mathcal{F}^*, P^*)$, namely there exists some event $F \in \mathcal{F}^*$ such that $P^*(F^c) = 0$ and

$$(\bar{v}^n, \bar{X}^n, \bar{Y}^n, \bar{\mathbf{B}}^n, \bar{\eta}^n, \tilde{\mathbf{B}}^n) \rightarrow (\bar{v}, \bar{X}, \bar{Y}, \bar{\mathbf{B}}, \bar{\eta}, \mathbf{0}) \quad \text{on } F.$$

Fix $\bar{\omega} \in F$. The rest of the argument will be made for such an $\bar{\omega}$ which will be suppressed from the notation. From (3.19) we have that as $n \rightarrow \infty$

$$r(\bar{X}^n(t)) = (\bar{X}_0^n(t))^+ + \sum_{k=1}^{\infty} k \bar{X}_k^n(t) \rightarrow (\bar{X}_0(t))^+ + \sum_{k=1}^{\infty} k \bar{X}_k(t) = r(\bar{X}(t))$$

uniformly in $t \in [0, T]$, and $r(\bar{X}(\cdot))$ is continuous and hence bounded. Let $\bar{\tau} \doteq \tau_{\bar{X}}$, where $\tau_{\bar{X}}$ is defined through (2.17), namely $\bar{\tau} = \inf\{t \in [0, T] : r(\bar{X}(t)) = 0\} \wedge T$. We will argue that (3.18) holds for all $t < \bar{\tau}$, $t = \bar{\tau}$ and $t > \bar{\tau}$.

For $t < \bar{\tau}$, we have $r(\bar{X}(t)) > 0$. Hence for each $k \in \mathbb{N}_0$,

$$(3.21) \quad 1_{[0, r_k(\bar{X}^n(s))]}(y) \rightarrow 1_{[0, r_k(\bar{X}(s))]}(y)$$

as $n \rightarrow \infty$ for λ_t -a.e. $(s, y) \in [0, t] \times [0, 1]$ since $\lambda_t\{(y, s) : y = r_k(\bar{X}(s))\} = 0$. From (3.21) and the uniform integrability of $(s, y) \mapsto (1_{[0, r_k(\bar{X}^n(s))]}(y) - 1_{[0, r_k(\bar{X}(s))]}(y))\varphi_k^n(s, y)$ (with respect to the normalized Lebesgue measure on $[0, T] \times [0, 1]$) which follows from (3.5) and the superlinearity of ℓ , we have that

$$\hat{B}_k^n(t) - \int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}(s))]}(y) \varphi_k^n(s, y) ds dy \rightarrow 0.$$

Also, from the bound in (3.5) it follows that

$$\int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}(s))]}(y) \varphi_k^n(s, y) ds dy \rightarrow \int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}(s))]}(y) \varphi_k(s, y) ds dy.$$

Combining the two convergence statements we have

$$(3.22) \quad \hat{B}_k^n(t) \rightarrow \int_{[0, t] \times [0, 1]} 1_{[0, r_k(\bar{X}(s))]}(y) \varphi_k(s, y) ds dy.$$

The above convergence along with (3.14) and (3.20) gives (3.18) for $t < \bar{\tau}$. Since (3.18) holds for $t < \bar{\tau}$, it also holds for $t = \bar{\tau}$ by continuity of $\bar{\mathbf{B}}$ and of the right side in (3.18).

Now suppose $T \geq t > \bar{\tau}$. Since $r(\bar{X}(\cdot))$ is continuous, we see from the definition of $\bar{\tau}$ that $r(\bar{X}(\bar{\tau})) = 0$. Noting that $r(\bar{X}^n(\cdot))$ is nonnegative and nonincreasing, so is $r(\bar{X}(\cdot))$. Therefore, $r(\bar{X}(t)) = 0$ and $\bar{X}(t) = \mathbf{0}$ for $\bar{\tau} \leq t \leq T$. From this we see that the right-hand side of (3.18) remains constant for $\bar{\tau} \leq t \leq T$ and it suffices to show that $\bar{B}(t) = \bar{B}(\bar{\tau})$ for $\bar{\tau} < t \leq T$. From (3.9) it follows that, for $k \in \mathbb{N}$,

$$(3.23) \quad \sup_{\bar{\tau} < t \leq T} |\bar{B}_k^n(t) - \bar{B}_k^n(\bar{\tau})| = \bar{B}_k^n(T) - \bar{B}_k^n(\bar{\tau}) = \bar{X}_k^n(\bar{\tau}) - \bar{X}_k^n(T) \leq \bar{X}_k^n(\bar{\tau}),$$

which converges to $\bar{X}_k(\bar{\tau}) = 0$ as $n \rightarrow \infty$. Hence $\bar{B}_k(t) = \bar{B}_k(\bar{\tau})$ for $\bar{\tau} < t \leq T$ and this gives (3.18) for each $k \in \mathbb{N}$. Next we show $\bar{B}_0(t) = \bar{B}_0(\bar{\tau})$ for $\bar{\tau} < t \leq T$. From (3.6) and (3.7),

$$(3.24) \quad \begin{aligned} & \sup_{\bar{\tau} < t \leq T} |\bar{B}_0^n(t) - \bar{B}_0^n(\bar{\tau})| \\ & \leq \sup_{\bar{\tau} < t \leq T} |\bar{X}_0^n(t) - \bar{X}_0^n(\bar{\tau})| + \sup_{\bar{\tau} < t \leq T} |\bar{\eta}^n(t) - \bar{\eta}^n(\bar{\tau})| \\ & \quad + \sum_{k=1}^{\infty} |k-2| \sup_{\bar{\tau} < t \leq T} |\bar{B}_k^n(t) - \bar{B}_k^n(\bar{\tau})|. \end{aligned}$$

Since $\bar{X}_0^n(t) \geq -\frac{1}{n}$, we have

$$\begin{aligned} \sup_{\bar{\tau} < t \leq T} |\bar{X}_0^n(t) - \bar{X}_0^n(\bar{\tau})| &\leq \sup_{\bar{\tau} < t \leq T} |\bar{X}_0^n(t)| + |\bar{X}_0^n(\bar{\tau})| \leq \sup_{\bar{\tau} < t \leq T} (\bar{X}_0^n(t))^+ + \frac{1}{n} + (\bar{X}_0^n(\bar{\tau}))^+ + \frac{1}{n} \\ &\leq \sup_{\bar{\tau} < t \leq T} r(\bar{X}^n(t)) + r(\bar{X}^n(\bar{\tau})) + \frac{2}{n} \leq 2r(\bar{X}^n(\bar{\tau})) + \frac{2}{n}, \end{aligned}$$

where the last line follows from the fact that $r(\bar{X}^n(t))$ is nonincreasing for $t \in [0, T]$. From (3.10) and (3.9) it follows that

$$\begin{aligned} &\sup_{\bar{\tau} < t \leq T} |\bar{\eta}^n(t) - \bar{\eta}^n(\bar{\tau})| \\ &= \sup_{\bar{\tau} < t \leq T} 2 \sum_{k=1}^{\infty} \frac{1}{n} \int_{(\bar{\tau}, t] \times [0, 1]} 1_{\{\bar{X}_0^n(u-) < 0\}} 1_{[0, r_k(\bar{X}^n(u-))]}(y) N_k^{n\varphi_k^n}(du dy) \\ &\leq \sup_{\bar{\tau} < t \leq T} 2 \sum_{k=1}^{\infty} \frac{1}{n} \int_{(\bar{\tau}, t] \times [0, 1]} 1_{[0, r_k(\bar{X}^n(u-))]}(y) N_k^{n\varphi_k^n}(du dy) \\ &= \sup_{\bar{\tau} < t \leq T} 2 \sum_{k=1}^{\infty} |\bar{B}_k^n(t) - \bar{B}_k^n(\bar{\tau})|. \end{aligned}$$

Combining the above two estimates with (3.24), we see that as $n \rightarrow \infty$,

$$\begin{aligned} (3.25) \quad &\sup_{\bar{\tau} < t \leq T} |\bar{B}_0^n(t) - \bar{B}_0^n(\bar{\tau})| \leq 2r(\bar{X}^n(\bar{\tau})) + \frac{2}{n} + \sup_{\bar{\tau} < t \leq T} \sum_{k=1}^{\infty} (k+4) |\bar{B}_k^n(t) - \bar{B}_k^n(\bar{\tau})| \\ &\leq 2r(\bar{X}^n(\bar{\tau})) + \frac{2}{n} + \sum_{k=1}^{\infty} (k+4) \bar{X}_k^n(\bar{\tau}) \leq 7r(\bar{X}^n(\bar{\tau})) + \frac{2}{n} \\ &\rightarrow 7r(\bar{X}(\bar{\tau})) = 0, \end{aligned}$$

where the second inequality follows from (3.23). Since we have proved (3.18) for all $t < \bar{\tau}$, $t = \bar{\tau}$ and $t > \bar{\tau}$, part (c) follows.

(d) From (3.17) and a well-known characterization of the solution of the Skorohod problem (see, e.g., [32], Section 3.6.C), it suffices to show that $\bar{\eta}(0) = 0$, $\bar{\eta}(t) \geq 0$, $\bar{\eta}(t)$ is non-decreasing for $t \in [0, T]$ and $\int_0^T \bar{X}_0(t) \bar{\eta}(dt) = 0$. Since $\bar{\eta}^n(0) = 0$, $\bar{\eta}^n(t) \geq 0$ and $\bar{\eta}^n(t)$ is nondecreasing for $t \in [0, T]$, so is $\bar{\eta}$. It remains to show $\int_0^T \bar{X}_0(t) \bar{\eta}(dt) = 0$. Note that $\bar{\eta}^n(t)$ increases only when $\bar{X}_0^n(t-) < 0$, namely $\bar{X}_0^n(t-) = -\frac{1}{n}$. Therefore,

$$\int_0^T \left(\bar{X}_0^n(t-) + \frac{1}{n} \right) \bar{\eta}^n(dt) = 0.$$

From this we have

$$\begin{aligned} (3.26) \quad &\left| \int_0^T \bar{X}_0(t) \bar{\eta}(dt) \right| = \left| \int_0^T \bar{X}_0(t) \bar{\eta}(dt) - \int_0^T \left(\bar{X}_0^n(t-) + \frac{1}{n} \right) \bar{\eta}^n(dt) \right| \\ &\leq \left| \int_0^T \bar{X}_0(t) \bar{\eta}(dt) - \int_0^T \bar{X}_0(t) \bar{\eta}^n(dt) \right| \\ &\quad + \int_0^T |\bar{X}_0(t) - \bar{X}_0^n(t-)| \bar{\eta}^n(dt) + \frac{\bar{\eta}^n(T)}{n}. \end{aligned}$$

Since both $\bar{\eta}^n$ and $\bar{\eta}$ are nondecreasing, we see that $\bar{\eta}^n \rightarrow \bar{\eta}$ as finite measures on $[0, T]$. Combining this with the fact that $\bar{X}_0 \in \mathbb{C}_b([0, T] : \mathbb{R})$, we get

$$\left| \int_0^T \bar{X}_0(t) \bar{\eta}(dt) - \int_0^T \bar{X}_0(t) \bar{\eta}^n(dt) \right| \rightarrow 0$$

as $n \rightarrow \infty$. Also from continuity of \bar{X}_0 , we have uniform convergence of \bar{X}_0^n to \bar{X}_0 and hence

$$\int_0^T |\bar{X}_0(t) - \bar{X}_0^n(t-)| \bar{\eta}^n(dt) + \frac{\bar{\eta}^n(T)}{n} \leq \left(\sup_{0 \leq t \leq T} |\bar{X}_0^n(t-) - \bar{X}_0(t)| + \frac{1}{n} \right) \bar{\eta}^n(T) \rightarrow 0$$

as $n \rightarrow \infty$. Combining these two convergence results with (3.26), we see that $\int_0^T \bar{X}_0(t) \times \bar{\eta}(dt) = 0$. This proves part (d) and completes the proof. \square

4. Laplace upper bound.

In this section we prove the Laplace upper bound (2.18).

From (3.1), for every $n \in \mathbb{N}$, we can choose $\tilde{\varphi}^n \doteq (\tilde{\varphi}_k^n)_{k \in \mathbb{N}_0} \in \bar{\mathcal{A}}_b$ such that

$$-\frac{1}{n} \log E e^{-nh(X^n, Y^n)} \geq E \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\tilde{\varphi}_k^n(s, y)) ds dy + h(\bar{X}^{n, \tilde{\varphi}^n}, \bar{Y}^{n, \tilde{\varphi}^n}) \right\} - \frac{1}{n},$$

where $(\bar{X}^{n, \tilde{\varphi}^n}, \bar{Y}^{n, \tilde{\varphi}^n})$ are defined by (3.2)–(3.4) by replacing φ^n with $\tilde{\varphi}^n$. Since $\|h\|_{\infty} < \infty$,

$$\begin{aligned} & \sup_{n \in \mathbb{N}} E \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\tilde{\varphi}_k^n(s, y)) ds dy \\ & \leq \sup_{n \in \mathbb{N}} \left[-\frac{1}{n} \log E e^{-nh(X^n, Y^n)} - Eh(\bar{X}^{n, \tilde{\varphi}^n}, \bar{Y}^{n, \tilde{\varphi}^n}) + \frac{1}{n} \right] \\ & \leq 2\|h\|_{\infty} + 1 \doteq M_h. \end{aligned}$$

Now we modify $\tilde{\varphi}^n$ so that the last inequality holds not in the sense of expectation, but rather almost surely, for a possibly larger constant [see (3.5)]. Fix $\sigma \in (0, 1)$ and define

$$\tilde{\tau}^n \doteq \inf \left\{ t \in [0, T] : \sum_{k=0}^{\infty} \int_{[0, t] \times [0, 1]} \ell(\tilde{\varphi}_k^n(s, y)) ds dy > 2M_h\|h\|_{\infty}/\sigma \right\} \wedge T.$$

For $k \in \mathbb{N}_0$, letting $\varphi_k^n(s, y) \doteq \tilde{\varphi}_k^n(s, y) 1_{\{s \leq \tilde{\tau}^n\}} + 1_{\{s > \tilde{\tau}^n\}}$, $(s, y) \in [0, T] \times [0, 1]$, we have $\varphi^n \doteq (\varphi_k^n)_{k \in \mathbb{N}_0} \in \bar{\mathcal{A}}_b$ since $\tilde{\tau}^n$ is an $\{\mathcal{F}_t\}$ -stopping time. Also

$$E \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy \leq E \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\tilde{\varphi}_k^n(s, y)) ds dy$$

and

$$\begin{aligned} P(\varphi^n \neq \tilde{\varphi}^n) & \leq P \left(\sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\tilde{\varphi}_k^n(s, y)) ds dy > 2M_h\|h\|_{\infty}/\sigma \right) \\ & \leq \frac{\sigma}{2M_h\|h\|_{\infty}} E \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\tilde{\varphi}_k^n(s, y)) ds dy \leq \frac{\sigma}{2\|h\|_{\infty}}. \end{aligned}$$

Letting $(\bar{X}^{n, \varphi^n}, \bar{Y}^{n, \varphi^n})$ be defined through (3.2)–(3.4) using φ^n , we have

$$|Eh(\bar{X}^{n, \varphi^n}, \bar{Y}^{n, \varphi^n}) - Eh(\bar{X}^{n, \tilde{\varphi}^n}, \bar{Y}^{n, \tilde{\varphi}^n})| \leq 2\|h\|_{\infty} P(\varphi^n \neq \tilde{\varphi}^n) \leq \sigma.$$

Hence we have

$$-\frac{1}{n} \log E e^{-nh(X^n, Y^n)} \geq E \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy + h(\bar{X}^n, \bar{Y}^n) \right\} - \frac{1}{n} - \sigma$$

and

$$(4.1) \quad \sup_{n \in \mathbb{N}} \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy \leq 2M_h \|h\|_{\infty}/\sigma \doteq K_0, \quad \text{a.s. } P.$$

Now we can complete the proof of the Laplace upper bound. Recall that $h \in \mathbb{C}_b(\mathcal{D}_{\infty} \times \mathcal{D})$. Write $(\bar{v}^n, \bar{X}^n, \bar{Y}^n) \doteq (\bar{v}^{n, \varphi^n}, \bar{X}^{n, \varphi^n}, \bar{Y}^{n, \varphi^n})$, where \bar{v}^{n, φ^n} is as defined in (3.12) using φ^n . Noting from (4.1) that (3.5) is satisfied with $M_0 = K_0$, we have from Lemma 3.4 that $\{(\bar{v}^n, \bar{X}^n, \bar{Y}^n)\}$ is tight. Assume without loss of generality that $(\bar{v}^n, \bar{X}^n, \bar{Y}^n)$ converges along the whole sequence weakly to $(\bar{v}, \bar{X}, \bar{Y})$, given on some probability space $(\Omega^*, \mathcal{F}^*, P^*)$. By Lemma 3.5 we have $(\bar{X}, \bar{Y}) \in \mathcal{C}_T$ and $\bar{v} = \bar{v}^{\varphi}$ for some $\varphi \in \mathcal{S}_T(\bar{X}, \bar{Y})$ a.s. P^* , where \bar{v}^{φ} is as defined in (3.12) using φ . Owing to the topology used for the measure component and the relation (3.12), Lemma A.1 in [11] (see also [13], Appendix A.4.3, Lemma A.11) implies the lower semicontinuity of the cost that is needed for the second inequality below. Using Fatou's lemma and the definition of I_T in (2.13),

$$\begin{aligned} & \liminf_{n \rightarrow \infty} -\frac{1}{n} \log E e^{-nh(X^n, Y^n)} \\ & \geq \liminf_{n \rightarrow \infty} E \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy + h(\bar{X}^n, \bar{Y}^n) - \frac{1}{n} - \sigma \right\} \\ & \geq E^* \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy + h(\bar{X}, \bar{Y}) \right\} - \sigma \\ & \geq \inf_{(\zeta, \psi) \in \mathcal{D}_{\infty} \times \mathcal{D}} \{I_T(\zeta, \psi) + h(\zeta, \psi)\} - \sigma. \end{aligned}$$

Since $\sigma \in (0, 1)$ is arbitrary, this completes the proof of the Laplace upper bound.

5. Laplace lower bound. In this section we prove the Laplace lower bound (2.19).

The following lemma, which shows unique solvability of the ODE (2.15) and (2.16) for controls φ in a suitable class, is key in the proof.

LEMMA 5.1. *Fix $\sigma \in (0, 1)$. Given $(\zeta, \psi) \in \mathcal{C}_T$ with $I_T(\zeta, \psi) < \infty$, there exists $\varphi^* \in \mathcal{S}_T(\zeta, \psi)$ such that:*

- (a) $\sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^*(s, y)) ds dy \leq I_T(\zeta, \psi) + \sigma$.
- (b) *If $(\tilde{\zeta}, \tilde{\psi})$ is another pair in \mathcal{C}_T such that $\varphi^* \in \mathcal{S}_T(\tilde{\zeta}, \tilde{\psi})$, then $(\tilde{\zeta}, \tilde{\psi}) = (\zeta, \psi)$.*

PROOF. Since $I_T(\zeta, \psi) < \infty$, we can choose some $\varphi \in \mathcal{S}_T(\zeta, \psi)$ such that

$$\sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy \leq I_T(\zeta, \psi) + \frac{\sigma}{2}.$$

Next we will modify φ to get the desired φ^* . For $k \in \mathbb{N}_0$, let

$$\rho_k(t) \doteq 1_{\{r_k(\zeta(t))=0\}} + \frac{\int_0^1 1_{[0, r_k(\zeta(t))]}(y) \varphi_k(t, y) dy}{r_k(\zeta(t))} 1_{\{r_k(\zeta(t)) \neq 0\}},$$

$$\tilde{\rho}_k(t, y) \doteq \rho_k(t) 1_{[0, r_k(\zeta(t))]}(y) + 1_{[r_k(\zeta(t)), 1]}(y).$$

Then

$$\int_{[0,t] \times [0,1]} 1_{[0,r_k(\xi(s))]}(y) \tilde{\varphi}_k(s, y) ds dy = \int_{[0,t] \times [0,1]} 1_{[0,r_k(\xi(s))]}(y) \varphi_k(s, y) ds dy$$

and hence $(\tilde{\varphi}_k)_{k \in \mathbb{N}_0} \in \mathcal{S}_T(\xi, \psi)$. Since ℓ is convex and nonnegative and $\ell(1) = 0$, we have

$$\begin{aligned} \int_{[0,T] \times [0,1]} \ell(\tilde{\varphi}_k(s, y)) ds dy &= \int_0^T \int_{[0,1]} 1_{\{r_k(\xi(s)) \neq 0\}} r_k(\xi(s)) \ell(\rho_k(s)) ds dy \\ &\leq \int_{[0,T] \times [0,1]} \ell(\varphi_k(s, y)) ds dy \end{aligned}$$

for each $k \in \mathbb{N}_0$. Therefore, we can assume without loss of generality (and abusing notation) that $\varphi_k(t, y) = \rho_k(t) 1_{[0,r_k(\xi(t))]}(y) + 1_{[r_k(\xi(t)), 1]}(y)$ for some $\rho_k(t) \in [0, \infty)$, for each $k \in \mathbb{N}_0$ and $(t, y) \in [0, T] \times [0, 1]$. Fix $\varepsilon \in (0, 1)$. We will shrink the support of φ to get the desired φ^* for sufficiently small ε . For $t \in [0, T]$, let

$$\varphi_k^\varepsilon(t, y) = \frac{\rho_k(t)}{1 - \varepsilon} 1_{[0, (1-\varepsilon)r_k(\xi(t))]}(y) + 1_{[(1+\varepsilon)r_k(\xi(t)), 1]}(y).$$

Clearly $\varphi^\varepsilon \in \mathcal{S}_T(\xi, \psi)$. Note that $\varphi_k^\varepsilon(t, y) = 0$ for $(1 - \varepsilon)r_k(\xi(t)) < y < (1 + \varepsilon)r_k(\xi(t))$, which will be a key when we prove uniqueness in part (b). Recall τ_ξ introduced in (2.17). Then

$$\begin{aligned} &\sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k^\varepsilon(t, y)) dt dy - \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k(t, y)) dt dy \\ &= \sum_{k=0}^{\infty} \int_0^{\tau_\xi} \left[(1 - \varepsilon)r_k(\xi(t)) \ell\left(\frac{\rho_k(t)}{1 - \varepsilon}\right) + 2\varepsilon r_k(\xi(t)) \ell(0) - r_k(\xi(t)) \ell(\rho_k(t)) \right] dt \\ &= \sum_{k=0}^{\infty} \int_0^{\tau_\xi} r_k(\xi(t)) \left[\left(\rho_k(t) \log\left(\frac{\rho_k(t)}{1 - \varepsilon}\right) - \rho_k(t) + 1 - \varepsilon \right) + 2\varepsilon \right. \\ &\quad \left. - (\rho_k(t) \log \rho_k(t) - \rho_k(t) + 1) \right] dt \\ &= \sum_{k=0}^{\infty} \int_0^{\tau_\xi} r_k(\xi(t)) \left[\rho_k(t) \log\left(\frac{1}{1 - \varepsilon}\right) + \varepsilon \right] dt. \end{aligned}$$

From Lemma 3.2(b) we have

$$\begin{aligned} &\sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k^\varepsilon(t, y)) dt dy - \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k(t, y)) dt dy \\ &\leq \sum_{k=0}^{\infty} \int_0^{\tau_\xi} r_k(\xi(t)) \left[(\ell(\rho_k(t)) + 2) \log\left(\frac{1}{1 - \varepsilon}\right) + \varepsilon \right] dt \\ &= \log\left(\frac{1}{1 - \varepsilon}\right) \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k(t, y)) dt dy + 2\tau_\xi \log\left(\frac{1}{1 - \varepsilon}\right) + \tau_\xi \varepsilon \\ &\leq \left(I_T(\xi, \psi) + \frac{\sigma}{2} \right) \log\left(\frac{1}{1 - \varepsilon}\right) + 2T \log\left(\frac{1}{1 - \varepsilon}\right) + T\varepsilon. \end{aligned}$$

Choosing ε small enough so that the last display is no larger than $\frac{\sigma}{2}$, we have

$$\sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k^\varepsilon(s, y)) ds dy \leq \sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k(s, y)) ds dy + \frac{\sigma}{2} \leq I_T(\xi, \psi) + \sigma.$$

Part (a) then holds with $\varphi^* = \varphi^\varepsilon$ for such an ε .

We now show that part (b) is satisfied with such a φ^* . Suppose that, in addition to (ξ, ψ) , there is another pair of $(\tilde{\xi}, \tilde{\psi})$ such that $(\tilde{\xi}, \tilde{\psi}) \in \mathcal{C}_T$ and $\varphi^* \in \mathcal{S}_T(\tilde{\xi}, \tilde{\psi})$. Let $\tau \doteq \inf\{t \in [0, T] : \xi(t) \neq \tilde{\xi}(t)\} \wedge T$. We claim that $\tau = T$. Once the claim is verified, it follows from continuity of ξ and $\tilde{\xi}$ that $\xi(t) = \tilde{\xi}(t)$ for all $t \in [0, T]$. Then from (2.15), $\psi = \tilde{\psi}$ proving part (b).

Now we prove the claim that $\tau = T$. We will argue via contradiction. Suppose that $\tau < T$. To complete the proof, it suffices to reach the following contradiction

$$(5.1) \quad \xi(t) = \tilde{\xi}(t), \quad t \in [\tau, \tau + \delta] \text{ for some } \delta > 0.$$

From definition of τ and (2.15) it follows that $(\xi(t), r(\xi(t)), \psi(t)) = (\tilde{\xi}(t), r(\tilde{\xi}(t)), \tilde{\psi}(t))$ for all $t < \tau$. From Remark 2.15(a) we have that $r(\xi(\cdot)), r(\tilde{\xi}(\cdot)) \in \mathcal{C}$. Then by continuity, $(\xi(t), r(\xi(t)), \psi(t)) = (\tilde{\xi}(t), r(\tilde{\xi}(t)), \tilde{\psi}(t))$ for all $t \leq \tau$. If $r(\xi(\tau)) = r(\tilde{\xi}(\tau)) = 0$, then from Remark 2.15(c) we have $\xi(t) = \tilde{\xi}(t) = \mathbf{0}$ for all $t \geq \tau$, which gives (5.1). Now we show (5.1) for the remaining case: $r(\xi(\tau)) = r(\tilde{\xi}(\tau)) > 0$. For this, note that by continuity of $r(\xi)$ and $r(\tilde{\xi})$, there exists some $\delta > 0$ such that for all $t \in [\tau, \tau + \delta]$,

$$(5.2) \quad r(\xi(t)) > 0, \quad r(\tilde{\xi}(t)) > 0, \quad \left| \frac{r(\xi(t))}{r(\tilde{\xi}(t))} - 1 \right| < \varepsilon,$$

where ε is as in part (a) and recall that $\varphi^* = \varphi^\varepsilon$. We will argue in two steps.

Step 1: We will prove that

$$(5.3) \quad \xi_k(t) = \tilde{\xi}_k(t) \text{ for all } t \in [\tau, \tau + \delta], k \in \mathbb{N}.$$

Suppose not, namely there exists $k \in \mathbb{N}$ such that $\tau_k \doteq \inf\{t \in [\tau, \tau + \delta] : \xi_k(t) \neq \tilde{\xi}_k(t)\} \wedge T$ satisfies $\tau \leq \tau_k < \tau + \delta$. By continuity, we have $\xi_k(t) = \tilde{\xi}_k(t)$ for $t \leq \tau_k$. We must have $\xi_k(\tau_k) = \tilde{\xi}_k(\tau_k) > 0$, since otherwise $\xi_k(\tau_k) = \tilde{\xi}_k(\tau_k) = 0$ and so from Remark 2.15(b) $\xi_k(t) = \tilde{\xi}_k(t) = 0$ for all $t \geq \tau_k$, which contradicts the definition of τ_k . From (5.2) it then follows that

$$r_k(\xi(\tau_k)) = \frac{k \xi_k(\tau_k)}{r(\xi(\tau_k))} > 0,$$

$$|r_k(\xi(\tau_k)) - r_k(\tilde{\xi}(\tau_k))| = \left| \frac{k \xi_k(\tau_k)}{r(\xi(\tau_k))} - \frac{k \tilde{\xi}_k(\tau_k)}{r(\tilde{\xi}(\tau_k))} \right| = \frac{k \xi_k(\tau_k)}{r(\xi(\tau_k))} \left| 1 - \frac{r(\xi(\tau_k))}{r(\tilde{\xi}(\tau_k))} \right| < \varepsilon r_k(\xi(\tau_k)).$$

Once more by continuity, there exists some $\delta_k > 0$ such that last two inequalities hold for $t \in [\tau_k, \tau_k + \delta_k]$, namely

$$r_k(\xi(t)) > 0, \quad (1 - \varepsilon)r_k(\xi(t)) < r_k(\tilde{\xi}(t)) < (1 + \varepsilon)r_k(\xi(t)).$$

From construction of φ^ε , we see that for $t \in [\tau_k, \tau_k + \delta_k]$,

$$\int_{(\tau_k, t] \times [0, 1]} 1_{[0, r_k(\tilde{\xi}(s))]}(y) \varphi_k^\varepsilon(s, y) ds dy = \int_{(\tau_k, t] \times [0, 1]} 1_{[0, r_k(\xi(s))]}(y) \varphi_k^\varepsilon(s, y) ds dy.$$

It then follows from (2.16) that $\xi_k(t) = \tilde{\xi}_k(t)$ for all $t \leq \tau_k + \delta_k$. This contradicts the definition of τ_k . Therefore, (5.3) must hold.

Step 2: We will prove that

$$(5.4) \quad \xi_0(t) = \tilde{\xi}_0(t) \quad \text{for all } t \in [\tau, \tau + \delta].$$

Let $\eta(t) \doteq \xi_0(t) - \psi(t)$ and $\tilde{\eta}(t) \doteq \tilde{\xi}_0(t) - \tilde{\psi}(t)$. From properties of the Skorokhod map Γ (see, e.g., [32], Section 3.6.C), we have that

$$(5.5) \quad \eta(0) = 0, \quad \eta(t) \text{ is nondecreasing and } \int_0^T \xi_0(t) \eta(dt) = 0,$$

$$(5.6) \quad \tilde{\eta}(0) = 0, \quad \tilde{\eta}(t) \text{ is nondecreasing and } \int_0^T \tilde{\xi}_0(t) \tilde{\eta}(dt) = 0.$$

Consider $[\zeta_0(t) - \tilde{\zeta}_0(t)]^2$. Since $\zeta_0, \psi, \tilde{\zeta}_0, \tilde{\psi}$ are absolutely continuous, we have for $t \in [\tau, \tau + \delta]$,

$$\begin{aligned} (5.7) \quad (\zeta_0(t) - \tilde{\zeta}_0(t))^2 &= (\zeta_0(\tau) - \tilde{\zeta}_0(\tau))^2 + 2 \int_{\tau}^t (\zeta_0(s) - \tilde{\zeta}_0(s))(\zeta_0'(s) - \tilde{\zeta}_0'(s)) ds \\ &= 2 \int_{\tau}^t (\zeta_0(s) - \tilde{\zeta}_0(s))(\psi'(s) - \tilde{\psi}'(s)) ds \\ &\quad + 2 \int_{\tau}^t (\zeta_0(s) - \tilde{\zeta}_0(s))(\eta_0'(s) - \tilde{\eta}_0'(s)) ds. \end{aligned}$$

From (2.15) and (2.16) we see that for $t \in [\tau, \tau + \delta]$,

$$\begin{aligned} \psi(t) &= \sum_{k=1}^{\infty} (k-2)(p_k - \zeta_k(t)) - 2 \int_{[0,t] \times [0,1]} 1_{[0,r_0(\zeta(s))]}(y) \varphi_0^{\varepsilon}(s, y) ds dy, \\ \tilde{\psi}(t) &= \sum_{k=1}^{\infty} (k-2)(p_k - \tilde{\zeta}_k(t)) - 2 \int_{[0,t] \times [0,1]} 1_{[0,r_0(\tilde{\zeta}(s))]}(y) \varphi_0^{\varepsilon}(s, y) ds dy. \end{aligned}$$

Taking the difference of these two displays and using (5.3), we have that for $t \in [\tau, \tau + \delta]$,

$$(5.8) \quad \psi(t) - \tilde{\psi}(t) = -2 \int_{[0,t] \times [0,1]} (1_{[0,r_0(\zeta(s))]}(y) - 1_{[0,r_0(\tilde{\zeta}(s))]}(y)) \varphi_0^{\varepsilon}(s, y) ds dy.$$

Since for each fixed $y \geq 0$ the function $x \mapsto \frac{x}{x+y}$ is nondecreasing on $(-y, \infty)$, we have from (5.3) and (5.2) that if for some $t \in [\tau, \tau + \delta]$, $\zeta_0(t) \geq \tilde{\zeta}_0(t)$, then

$$r_0(\zeta(t)) = \frac{\zeta_0(t)}{\zeta_0(t) + \sum_{k=1}^{\infty} k \zeta_k(t)} = \frac{\zeta_0(t)}{\zeta_0(t) + \sum_{k=1}^{\infty} k \tilde{\zeta}_k(t)} \geq \frac{\tilde{\zeta}_0(t)}{\zeta_0(t) + \sum_{k=1}^{\infty} k \tilde{\zeta}_k(t)} = r_0(\tilde{\zeta}(t)).$$

Therefore, for $t \in [\tau, \tau + \delta]$,

$$1_{[0,r_0(\zeta(t))]}(y) \geq 1_{[0,r_0(\tilde{\zeta}(t))]}(y) \quad \text{when } \zeta_0(t) \geq \tilde{\zeta}_0(t)$$

and similarly

$$1_{[0,r_0(\zeta(t))]}(y) \leq 1_{[0,r_0(\tilde{\zeta}(t))]}(y) \quad \text{when } \zeta_0(t) \leq \tilde{\zeta}_0(t).$$

Combining these two inequalities with (5.8), we see that

$$(5.9) \quad (\zeta_0(s) - \tilde{\zeta}_0(s))(\psi'(s) - \tilde{\psi}'(s)) \leq 0, \quad \text{a.e. } s \in [\tau, \tau + \delta].$$

Next from (5.5) and (5.6) we see that for $t \in [\tau, \tau + \delta]$,

$$\begin{aligned} &\int_{\tau}^t 1_{\{\zeta_0(s) > \tilde{\zeta}_0(s)\}} (\zeta_0(s) - \tilde{\zeta}_0(s))(\eta_0'(s) - \tilde{\eta}_0'(s)) ds \\ &\leq \int_{\tau}^t 1_{\{\zeta_0(s) > \tilde{\zeta}_0(s)\}} (\zeta_0(s) - \tilde{\zeta}_0(s)) \eta_0'(s) ds \\ &\leq \int_{\tau}^t 1_{\{\zeta_0(s) > 0\}} \zeta_0(s) \eta_0(ds) = 0, \end{aligned}$$

and similarly

$$\int_{\tau}^t 1_{\{\zeta_0(s) < \tilde{\zeta}_0(s)\}} (\zeta_0(s) - \tilde{\zeta}_0(s))(\eta_0'(s) - \tilde{\eta}_0'(s)) ds \leq 0.$$

Combining these two inequalities with (5.9) and (5.7), we have for $t \in [\tau, \tau + \delta]$, $[\zeta_0(t) - \tilde{\zeta}_0(t)]^2 \leq 0$, proving (5.4). Combining (5.3) and (5.4) gives (5.1) and completes the proof. \square

We can now complete the proof of the Laplace lower bound. Fix $h \in \mathbb{C}_b(\mathcal{D}_\infty \times \mathcal{D})$ and $\sigma \in (0, 1)$. Fix some σ -optimal $(\zeta^*, \psi^*) \in \mathcal{C}_T$ with $I_T(\zeta^*, \psi^*) < \infty$, namely

$$I_T(\zeta^*, \psi^*) + h(\zeta^*, \psi^*) \leq \inf_{(\zeta, \psi) \in \mathcal{D}_\infty \times \mathcal{D}} \{I_T(\zeta, \psi) + h(\zeta, \psi)\} + \sigma.$$

Let $\varphi^* \in \mathcal{S}_T(\zeta^*, \psi^*)$ be as in Lemma 5.1 (with (ζ, ψ) there replaced by (ζ^*, ψ^*)). For each $n \in \mathbb{N}$ and $(s, y) \in [0, T] \times [0, 1]$, consider the deterministic control

$$\begin{aligned} \varphi_k^n(s, y) &\doteq \frac{1}{n} 1_{\{\varphi_k^*(s, y) \leq \frac{1}{n}\}} + \varphi_k^*(s, y) 1_{\{\frac{1}{n} < \varphi_k^*(s, y) < n\}} + n 1_{\{\varphi_k^*(s, y) \geq n\}}, \quad k \leq n, \\ \varphi_k^n(s, y) &\doteq 1, \quad k > n. \end{aligned}$$

Then $\varphi^n \doteq (\varphi_k^n) \in \bar{\mathcal{A}}_b$ and from (3.1) we have

$$-\frac{1}{n} \log E e^{-nh(X^n, Y^n)} \leq E \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy + h(\bar{X}^n, \bar{Y}^n) \right\},$$

where (\bar{X}^n, \bar{Y}^n) are given as in (3.2)–(3.4). Noting that for all $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ and $(s, y) \in [0, T] \times [0, 1]$, $\ell(\varphi_k^n(s, y)) \leq \ell(\varphi_k^*(s, y))$, we have from Lemma 5.1(a) that (3.5) holds with M_0 replaced by $I_T(\zeta^*, \psi^*) + 1$. Define $\{\bar{v}^n\}$ as in (3.12) with controls φ^n . From Lemma 3.4 it follows that $\{(\bar{v}^n, \bar{X}^n, \bar{Y}^n)\}$ is tight. Assume without loss of generality that $(\bar{v}^n, \bar{X}^n, \bar{Y}^n)$ converges along the whole sequence weakly to $(\bar{v}, \bar{X}, \bar{Y})$, given on some probability space $(\Omega^*, \mathcal{F}^*, P^*)$. From the construction of φ^n we must have $\bar{v} = \bar{v}^{\varphi^*}$ a.s. P^* , where \bar{v}^{φ^*} is as defined in (3.12) using φ^* . By Lemma 3.5 we have $(\bar{X}, \bar{Y}) \in \mathcal{C}_T$ and $\varphi^* \in \mathcal{S}_T(\bar{X}, \bar{Y})$ a.s. P^* . From Lemma 5.1(b) it now follows that $(\bar{X}, \bar{Y}) = (\zeta^*, \psi^*)$ a.s. P^* . Finally, from Lemma 5.1(a),

$$\begin{aligned} \limsup_{n \rightarrow \infty} -\frac{1}{n} \log E e^{-nh(X^n, Y^n)} &\leq \limsup_{n \rightarrow \infty} E \left\{ \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy + h(\bar{X}^n, \bar{Y}^n) \right\} \\ &\leq \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^*(s, y)) ds dy + E^* h(\bar{X}, \bar{Y}) \\ &= \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^*(s, y)) ds dy + h(\zeta^*, \psi^*) \\ &\leq I_T(\zeta^*, \psi^*) + h(\zeta^*, \psi^*) + \sigma \\ &\leq \inf_{(\zeta, \psi) \in \mathcal{D}_\infty \times \mathcal{D}} \{I_T(\zeta, \psi) + h(\zeta, \psi)\} + 2\sigma. \end{aligned}$$

Since $\sigma \in (0, 1)$ is arbitrary, this completes the proof of the Laplace lower bound.

6. Compact sublevel sets. In this section we prove that the function I_T defined in (2.13) is a rate function, namely the set $\Gamma_N \doteq \{(\zeta, \psi) \in \mathcal{D}_\infty \times \mathcal{D} : I_T(\zeta, \psi) \leq N\}$ is compact for each fixed $N \in [0, \infty)$. Since the proof (as is usual) is very similar to the proof of the Laplace upper bound we will only provide details on steps that are significantly different.

Take any sequence $\{(\zeta^n, \psi^n)_{n \in \mathbb{N}}\} \subset \Gamma_N$. Then $(\zeta^n, \psi^n) \in \mathcal{C}_T$ and there exists some $\frac{1}{n}$ -optimal $\varphi^n \in \mathcal{S}_T(\zeta^n, \psi^n)$, namely

$$(6.1) \quad \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^n(s, y)) ds dy \leq I_T(\zeta^n, \psi^n) + \frac{1}{n} \leq N + \frac{1}{n}.$$

Recalling (2.15) and (2.16) and letting $\eta^n(t) \doteq \zeta_0^n(t) - \psi^n(t)$, we can write for $t \in [0, T]$,

$$(6.2) \quad \zeta_0^n(t) = \Gamma(\psi^n)(t) = \psi^n(t) + \eta^n(t) = \sum_{k=0}^{\infty} (k-2) B_k^n(t) + \eta^n(t),$$

where

$$(6.3) \quad B_k^n(t) \doteq \int_{[0,t] \times [0,1]} 1_{[0,r_k(\zeta^n(s))]}(y) \varphi_k^n(s, y) ds dy, \quad k \in \mathbb{N}_0.$$

From standard properties of the one-dimensional Skorokhod problem we have

$$(6.4) \quad \eta^n(0) = 0, \eta^n(t) \text{ is nondecreasing and } \int_0^T 1_{\{\zeta_0^n(t) > 0\}} \eta^n(dt) = 0.$$

Write $\mathbf{B}^n = (B_k^n)_{n \in \mathbb{N}_0}$ and let \mathbf{v}^n be defined as in (3.12) with deterministic controls φ^n . The following lemma shows that $\{(\mathbf{v}^n, \zeta^n, \psi^n, \mathbf{B}^n, \eta^n)\}$ is pre-compact. The proof is similar to that of Lemma 3.4 and is therefore omitted.

LEMMA 6.1. $\{(\mathbf{v}^n, \zeta^n, \psi^n, \mathbf{B}^n, \eta^n)\}$ is pre-compact in $[\mathcal{M}_{FC}([0, T] \times [0, 1])]^\infty \times \mathcal{C}_\infty \times \mathcal{C} \times \mathcal{C}_\infty \times \mathcal{C}$.

The following lemma characterizes limit points of $(\mathbf{v}^n, \zeta^n, \psi^n, \mathbf{B}^n, \eta^n)$.

LEMMA 6.2. Suppose $(\mathbf{v}^n, \zeta^n, \psi^n, \mathbf{B}^n, \eta^n)$ converges along a subsequence to $(\mathbf{v}, \zeta, \psi, \mathbf{B}, \eta) \in [\mathcal{M}([0, T] \times [0, 1])]^\infty \times \mathcal{C}_\infty \times \mathcal{C} \times \mathcal{C}_\infty \times \mathcal{C}$. Then the following hold.

(a) For each $k \in \mathbb{N}_0$, $v_k \ll \lambda_T$, and letting $\varphi_k \doteq \frac{dv_k}{d\lambda_T}$, $\sum_{k=0}^{\infty} \int_{[0,T] \times [0,1]} \ell(\varphi_k(s, y)) ds dy \leq N$.

(b) For each $t \in [0, T]$,

$$\zeta_0(t) = \Gamma(\psi)(t) = \psi(t) + \eta(t), \quad \psi(t) = \sum_{k=0}^{\infty} (k-2) B_k(t)$$

$$\zeta_k(t) = p_k - B_k(t), \quad k \in \mathbb{N}.$$

(c) For each $t \in [0, T]$,

$$(6.5) \quad B_k(t) = \int_{[0,t] \times [0,1]} 1_{[0,r_k(\zeta(s))]}(y) \varphi_k(s, y) ds dy, \quad k \in \mathbb{N}_0,$$

and in particular $(\zeta, \psi) \in \mathcal{C}_T$ and $\varphi \in \mathcal{S}_T(\zeta, \psi)$.

PROOF. Assume without loss of generality that

$$(6.6) \quad (\mathbf{v}^n, \zeta^n, \psi^n, \mathbf{B}^n, \eta^n) \rightarrow (\mathbf{v}, \zeta, \psi, \mathbf{B}, \eta)$$

as $n \rightarrow \infty$ along the whole sequence. Much of the proof is similar to that of Lemma 3.5 except the proof of (6.5) for $k = 0$. Thus we only give details for the latter statement.

From (6.6) and arguments similar to Lemma 3.3 it follows that

$$r(\zeta^n(t)) = (\zeta_0^n(t))^+ + \sum_{k=1}^{\infty} k \zeta_k^n(t) \rightarrow (\zeta_0(t))^+ + \sum_{k=1}^{\infty} k \zeta_k(t) = r(\zeta(t))$$

uniformly in $t \in [0, T]$ as $n \rightarrow \infty$. Therefore, $r(\zeta(\cdot))$ is continuous. Let $\tau \doteq \inf\{t \in [0, T] : r(\zeta(t)) = 0\} \wedge T$. We will argue that (6.5), for $k = 0$, holds for all $t < \tau$, $t = \tau$ and $t > \tau$. The proof of the cases $t < \tau$ and $t = \tau$ is similar to that of (3.18) and is omitted.

Now consider $T \geq t > \tau$. From (6.4) and (6.2), for $\tau < t \leq T$,

$$\begin{aligned} |\eta^n(t) - \eta^n(\tau)| &= \int_{\tau}^t d\eta^n(s) = \int_{\tau}^t 1_{\{\zeta_0^n(s)=0\}} d\eta^n(s) \\ &= \int_{\tau}^t 1_{\{\zeta_0^n(s)=0\}} \left(d\zeta_0^n(s) - \sum_{k=0}^{\infty} (k-2) dB_k^n(s) \right). \end{aligned}$$

From (6.3) we see that $\int_{\tau}^t 1_{\{\zeta_0^n(s)=0\}} dB_0^n(s) = 0$. Also since ζ_0^n is nonnegative and absolutely continuous, we have $1_{\{\zeta_0^n(s)=0\}}(\zeta_0^n)'(s) = 0$ for a.e. $s \in [0, T]$. Therefore,

$$|\eta^n(t) - \eta^n(\tau)| \leq \sum_{k=1}^{\infty} |k-2| |B_k^n(t) - B_k^n(\tau)|.$$

Applying the triangle inequality to (6.2) and using this estimate, we see that

$$\sup_{\tau < t \leq T} |B_0^n(t) - B_0^n(\tau)| \leq \sup_{\tau < t \leq T} |\zeta_0^n(t) - \zeta_0^n(\tau)| + 2 \sum_{k=1}^{\infty} |k-2| \sup_{\tau < t \leq T} |B_k^n(t) - B_k^n(\tau)|.$$

Now as in the proof of (3.25) we have $\sup_{\tau < t \leq T} |B_0^n(t) - B_0^n(\tau)| \leq 4r(\zeta^n(\tau))$, which converges to $4r(\zeta(\tau)) = 0$ as $n \rightarrow \infty$. Hence $B_0(t) = B_0(\tau)$ for $\tau < t \leq T$ and this gives (6.5) for $k = 0$.

Since we have proved (6.5) for $k = 0$ and all $t < \tau$, $t = \tau$ and $t > \tau$, the proof is complete. \square

PROOF OF COMPACT SUBLVEL SETS Γ_M . Now we are ready to prove that Γ_M is compact for each fixed $M \in [0, \infty)$. Recall (ζ^n, ψ^n) introduced above (6.1) and ν^n introduced above Lemma 6.1. From Lemma 6.1 we have pre-compactness of $\{(\nu^n, \zeta^n, \psi^n)\}$ in $[\mathcal{M}([0, T] \times [0, 1])]^\infty \times \mathcal{C}_\infty \times \mathcal{C}$. Assume without loss of generality that (ν^n, ζ^n, ψ^n) converges along the whole sequence to some (ν, ζ, ψ) . By Lemma 6.2 $(\zeta, \psi) \in \mathcal{C}_T$ and $\nu = \nu^\varphi$, where for $k \in \mathbb{N}_0$, ν_k^φ is as defined by the right side of (3.12) replacing φ_k^n with φ_k , and

$$I_T(\zeta, \psi) \leq \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy \leq M.$$

Therefore, $(\zeta, \psi) \in \Gamma_M$ which proves that Γ_M is compact. \square

REMARK 6.3. Suppose that for all $n \in \mathbb{N}$, $(\zeta^n, \psi^n) = (\zeta, \psi)$ for some $(\zeta, \psi) \in \mathcal{C}_T$ with $I_T(\zeta, \psi) < \infty$ and $M = I_T(\zeta, \psi)$. Then taking φ^n satisfying (6.1) (with (ζ^n, ψ^n) replaced with (ζ, ψ)), we see from the above argument that there exists some $\varphi \in \mathcal{S}_T(\zeta, \psi)$ such that

$$I_T(\zeta, \psi) \leq \sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy \leq I_T(\zeta, \psi),$$

namely $I_T(\zeta, \psi)$ is achieved at some $\varphi \in \mathcal{S}_T(\zeta, \psi)$.

7. Calculus of variations problem. In this section we study a calculus of variations problem that is key in proof of Theorem 2.6. We begin by giving an overview of the proof strategy. Let $0 \leq \mathbf{q} \leq \mathbf{p}$. First note that, in view of Remark 2.14 and since, as noted in Section 2.4, $\{(nX_0^n(\sigma_j^n) + 1, nX_k^n(\sigma_j^n)), k, j \in \mathbb{N}\}$ has the same distribution as $\{A(j), V_k(j), k, j \in \mathbb{N}\}$, where $\{\sigma_j^n\}$ denote the jump instants of the process X^n , the set

$E^{n,\varepsilon}(\mathbf{q})$ can be written, in distributionally equivalent form (namely the probabilities of the events on the left and the right of the display below are the same), as

$$(7.1) \quad \begin{aligned} E^{n,\varepsilon}(\mathbf{q}) &= \left\{ \exists t_1, t_2 \in [0, \infty) \text{ such that } X_0^n(t_1-) = X_0^n(t_2) = -1/n, \right. \\ &\quad X_0^n(t) > -1/n \text{ for } t \in [t_1, t_2], \\ &\quad \left. |X_k^n(t_1-) - X_k^n(t_2) - q_k| \leq \varepsilon \text{ for all } k \in \mathbb{N} \right\}. \end{aligned}$$

Here t_1 (resp. t_2) corresponds to the time instant the first vertex (resp. the last edge) in a component is woken up (resp. is formed).

For $t_2 \geq t_1 \geq 0$ and $(\zeta, \psi) \in \mathcal{C}_{t_2}$, define

$$(7.2) \quad I_{t_1, t_2}(\zeta, \psi) \doteq \inf_{\varphi \in \mathcal{S}_{t_2}(\zeta, \psi)} \sum_{k=0}^{\infty} \int_{[t_1, t_2] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy.$$

Further, for $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \mathbb{R}_+^\infty$, define

$$\begin{aligned} \mathcal{J}_{t_1, t_2}^0(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) &\doteq \{(\zeta, \psi) \in \mathcal{C}_{t_2} : \zeta(t_1) = \mathbf{x}^{(1)}, \zeta(t_2) = \mathbf{x}^{(2)}\}, \\ \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) &\doteq \{(\zeta, \psi) \in \mathcal{J}_{t_1, t_2}^0(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) : \psi(t) \geq \psi(t_1) - x_0^{(1)} \text{ for } t \in (t_1, t_2)\}, \\ \mathcal{J}_{t_1, t_2}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) &\doteq \{(\zeta, \psi) \in \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) : dr(\zeta(t))/dt = -2 \text{ for a.e. } t \in (t_1, t_2)\}, \end{aligned}$$

and

$$(7.3) \quad I_{t_1, t_2}^j(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \doteq \inf_{(\zeta, \psi) \in \mathcal{J}_{t_1, t_2}^j(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})} I_{t_1, t_2}(\zeta, \psi), \quad j = 0, 1, 2.$$

Here as usual, the infimum over an empty set is infinity.

The proof of Theorem 2.6 proceeds through the following steps. Let $\tau \doteq \frac{1}{2} \sum_{k=1}^{\infty} k q_k$ and assume $\sum_{k=1}^{\infty} k q_k > 2 \sum_{k=1}^{\infty} q_k$. Note that the limit as $\varepsilon \rightarrow 0$ in fact exists because the set $E^{n,\varepsilon}(\mathbf{q})$ is decreasing as ε decreases.

- Lemma 8.4 shows the lower bound

$$(7.4) \quad \liminf_{\varepsilon \rightarrow 0} \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \geq -I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})).$$

- In Lemma 8.2 we show the upper bound

$$(7.5) \quad \begin{aligned} &\limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \\ &\leq - \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0, t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1, t_1 + \tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))]. \end{aligned}$$

- Lemma 8.3 shows that when $p_1 = 0$ the upper and lower bounds coincide.

- Finally Proposition 7.7 shows that

$$I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})) = H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}) + K(\mathbf{q})$$

completing the proof of Theorem 2.6.

Note that for $(\zeta, \psi) \in \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$, $\zeta_0(t) = x_0^{(1)} + \psi(t) - \psi(t_1)$ for $t \in [t_1, t_2]$. Intuitively, on the event $\{(X^n, Y^n) \in \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})\}$ the exploration remains in the same component over $[t_1, t_2]$, and on the smaller event $\{(X^n, Y^n) \in \mathcal{J}_{t_1, t_2}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})\}$ the exploration pace matches that for the discrete-time exploration process (with time steps of length $1/n$), in which at each step 2 half-edges are killed. The main idea in the proof of the theorem is that in characterizing the asymptotics of the probability of interest one can restrict to

$\mathcal{J}_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q}))$, which roughly means that one can restrict to trajectories that avoid the boundary and whose evolution matches that of the original discrete time process of interest removing the artificial “continuous time” aspect of the evolution.

Define for $\mathbf{x} = (x_k)_{k \in \mathbb{N}_0} \in \mathbb{R}_+^\infty$ and $\boldsymbol{\beta} = (\beta_k)_{k \in \mathbb{N}_0} \in \mathbb{R} \times [-1, 0]^\infty$ with $\sum_{k=1}^\infty \beta_k \geq -1$,

$$(7.6) \quad L(\mathbf{x}, \boldsymbol{\beta}) \doteq \sum_{k=0}^\infty \nu(k|\boldsymbol{\beta}) \log \left(\frac{\nu(k|\boldsymbol{\beta})}{\mu(k|\mathbf{x})} \right), \quad L_k(\mathbf{x}, \boldsymbol{\beta}) \doteq \nu(k|\boldsymbol{\beta}) \log \left(\frac{\nu(k|\boldsymbol{\beta})}{\mu(k|\mathbf{x})} \right),$$

where

$$(7.7) \quad \nu(0|\boldsymbol{\beta}) \doteq 1 + \sum_{k=1}^\infty \beta_k, \quad \nu(k|\boldsymbol{\beta}) \doteq -\beta_k, \quad k \in \mathbb{N},$$

$$(7.8) \quad \mu(k|\mathbf{x}) \doteq r_k(\mathbf{x}), \mathbf{x} \neq \mathbf{0}, \quad \mu(k|\mathbf{x}) \doteq 1_{\{k=0\}}, \quad \mathbf{x} = \mathbf{0}, k \in \mathbb{N}_0.$$

We set $L(\mathbf{x}, \boldsymbol{\beta}) = \infty$, if $\boldsymbol{\beta} \in \mathbb{R} \times [-1, 0]^\infty$ and $\sum_{k=1}^\infty \beta_k < -1$. Note that β_0 actually does not play a role in the definition of $L(\mathbf{x}, \boldsymbol{\beta})$ or $\nu(\cdot|\boldsymbol{\beta})$. Later on $(\mathbf{x}, \boldsymbol{\beta})$ will be usually replaced by $(\boldsymbol{\xi}(t), \boldsymbol{\xi}'(t))$ for some absolutely continuous path $\boldsymbol{\xi} = (\xi_k)_{k \in \mathbb{N}_0}$ and $t \geq 0$, where $\boldsymbol{\xi}'(t) \doteq (\xi'_k(t))_{k \in \mathbb{N}_0}$.

In the next six lemmas $\mathbf{x}^{(1)} \doteq (x_0^{(1)}, \mathbf{p}^{(1)})$ and $\mathbf{x}^{(2)} \doteq (x_0^{(2)}, \mathbf{p}^{(2)})$ where $x_0^{(1)}, x_0^{(2)} \in \mathbb{R}_+$ and $\mathbf{0} \leq \mathbf{p}^{(2)} \leq \mathbf{p}^{(1)} \leq \mathbf{p}$. Let $\zeta \doteq \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$. Define

$$(7.9) \quad \zeta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \doteq \frac{1}{2}(r(\mathbf{x}^{(1)}) - r(\mathbf{x}^{(2)})) = \frac{1}{2} \left((x_0^{(1)} - x_0^{(2)}) + \sum_{k=1}^\infty k(p_k^{(1)} - p_k^{(2)}) \right).$$

We write $\zeta \equiv \zeta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ for short in the next six lemmas. The following lemma relates I^1 , I^2 and L .

LEMMA 7.1. Fix $t_1 \in [0, \infty)$. Suppose $\zeta \geq 0$. Let $\mathbf{x}^{(0)} \doteq (0, \mathbf{p})$. Suppose there exists some $(\boldsymbol{\xi}^*, \psi^*) \in \mathcal{J}_{0,t_1}^0(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})$ such that $I_{0,t_1}(\boldsymbol{\xi}^*, \psi^*) < \infty$. Then

$$(7.10) \quad \inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = I_{t_1, t_1 + \zeta}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}).$$

Furthermore, for $(\boldsymbol{\xi}, \psi) \in \mathcal{J}_{t_1, t_1 + \zeta}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$,

$$(7.11) \quad I_{t_1, t_1 + \zeta}(\boldsymbol{\xi}, \psi) = \int_{t_1}^{t_1 + \zeta} L(\boldsymbol{\xi}(s), \boldsymbol{\xi}'(s)) ds,$$

and if $I_{t_1, t_1 + \zeta}(\boldsymbol{\xi}, \psi) < \infty$, then $\sum_{k=1}^\infty \zeta'_k(t) \geq -1$ for a.e. $t \in [t_1, t_1 + \zeta]$. In particular,

$$(7.12) \quad \begin{aligned} \inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) &= I_{t_1, t_1 + \zeta}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \\ &= \inf_{(\boldsymbol{\xi}, \psi) \in \mathcal{J}_{t_1, t_1 + \zeta}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})} \int_{t_1}^{t_1 + \zeta} L(\boldsymbol{\xi}(s), \boldsymbol{\xi}'(s)) ds. \end{aligned}$$

LEMMA 7.2. Suppose that $\sum_{k=1}^\infty k z_k + z_0 > 2 \sum_{k=1}^\infty z_k$ and, $x_0^{(2)} > 0$ or $z_1 > 0$. Then there is a unique $\beta \equiv \beta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in (0, 1)$ such that

$$(7.13) \quad \sum_{k=1}^\infty k z_k = (1 - \beta^2) \sum_{k=1}^\infty \frac{k z_k}{1 - \beta^k} + x_0^{(2)} - \beta^2 x_0^{(1)}.$$

The construction given below will be used to give an explicit representation for the minimizer for the right side of (7.12).

CONSTRUCTION 1. *Suppose that either (i) or (ii) holds, where:*

- (i) $x_0^{(2)} = 0$ and $z_1 = 0$.
- (ii) $\sum_{k=1}^{\infty} kz_k + z_0 > 2 \sum_{k=1}^{\infty} z_k$ and, $x_0^{(2)} > 0$ or $z_1 > 0$.

Let $\beta \equiv \beta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in [0, 1)$ be 0 in case (i) and be the unique solution in $(0, 1)$ of (7.13) in case (ii) (as ensured by Lemma 7.2). Note that β satisfies (7.13) in both cases (i) and (ii).

Define ς as in (7.9) and suppose that $\varsigma \geq 0$. Let $\tilde{\varsigma} \doteq \varsigma/(1 - \beta^2)$ and $\tilde{z}_k \doteq z_k/(1 - \beta^k)$ for $k \in \mathbb{N}$. Fix $t_1 \geq 0$ and let $\mathbf{x}^{(0)}, (\xi^*, \psi^*)$ be as in Lemma 7.1. Define $(\tilde{\xi}, \tilde{\psi})$ by $(\tilde{\xi}(t), \tilde{\psi}(t)) = (\xi^*(t), \psi^*(t))$ for $t \in [0, t_1]$ and for $t \in [t_1, t_1 + \varsigma]$

$$(7.14) \quad \tilde{\xi}_k(t) \doteq p_k^{(1)} - \tilde{z}_k \left[1 - \left(1 - \frac{t - t_1}{\tilde{\varsigma}} \right)^{k/2} \right], \quad k \in \mathbb{N},$$

$$(7.15) \quad \tilde{\xi}_0(t) \doteq x_0^{(1)} + \sum_{k=1}^{\infty} k(p_k^{(1)} - \tilde{\xi}_k(t)) - 2(t - t_1),$$

$$(7.16) \quad \tilde{\psi}(t) \doteq \tilde{\psi}(t_1) + \sum_{k=1}^{\infty} k(p_k^{(1)} - \tilde{\xi}_k(t)) - 2(t - t_1).$$

The next two lemmas give some properties of the various quantities in the above construction. Let

$$\Xi \doteq \left\{ (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) : \text{for } i = 1, 2, \mathbf{x}^{(i)} \doteq (x_0^{(i)}, \mathbf{p}^{(i)}), x_0^{(i)} \in \mathbb{R}_+, \right. \\ \left. \mathbf{0} \leq \mathbf{p}^{(2)} \leq \mathbf{p}^{(1)} \leq \mathbf{p} \text{ and } \sum_{k=1}^{\infty} k(p_k^{(1)} - p_k^{(2)}) + (x_0^{(1)} - x_0^{(2)}) > 2 \sum_{k=1}^{\infty} (p_k^{(1)} - p_k^{(2)}) \right\}.$$

We will equip Ξ with the topology corresponding to coordinatewise convergence.

LEMMA 7.3. *Both β and $x_0^{(2)} \log \beta$ are continuous on Ξ : for $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \in \Xi$ with $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \rightarrow (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \Xi$, $\beta^n \doteq \beta(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \rightarrow \beta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \doteq \beta$ and $x_0^{(2),n} \log \beta^n \rightarrow x_0^{(2)} \log \beta$.*

LEMMA 7.4. *Suppose that $\varsigma \geq 0$. Also suppose that $\sum_{k=1}^{\infty} kz_k + z_0 > 2 \sum_{k=1}^{\infty} z_k$. Fix $t_1 \geq 0$. Let (ξ^*, ψ^*) be as in Lemma 7.1 and $(\tilde{\xi}, \tilde{\psi})$ be as introduced in Construction 1. Then:*

- (a) $\varsigma \leq \tilde{\varsigma} = \frac{1}{2}(x_0^{(1)} + \sum_{k=1}^{\infty} k\tilde{z}_k)$.
- (b) $(\tilde{\xi}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_1 + \varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$.
- (c) $\tilde{\xi}_0(t) > 0$ for $t \in (t_1, t_1 + \varsigma)$.

The next lemma calculates $\int_{t_1}^{t_1 + \varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds$ for $(\tilde{\xi}, \tilde{\psi})$ introduced in Construction 1. Recall that $\mathbf{z} = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$.

LEMMA 7.5. *Suppose that $\varsigma \geq 0$. Suppose that either (i) or (ii) in Construction 1 is satisfied. Also, let (ξ^*, ψ^*) be as in Lemma 7.1 with some $t_1 \geq 0$ and let $(\tilde{\xi}, \tilde{\psi})$ be given as in Construction 1. Define the function $\tilde{K}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ by*

$$\tilde{K}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \doteq \frac{z_0 + \sum_{k=1}^{\infty} kz_k}{2} \log(1 - \beta^2) - \sum_{k=1}^{\infty} z_k \log(1 - \beta^k) + x_0^{(2)} \log \beta.$$

For $\mathbf{x} \in \mathbb{R} \times \mathbb{R}_+^\infty$ such that $x_0 + \sum_{k=1}^\infty kx_k \geq 0$, define $\tilde{H}(\mathbf{x})$ by

$$\tilde{H}(\mathbf{x}) \doteq \sum_{k=1}^\infty x_k \log x_k - \frac{x_0 + \sum_{k=1}^\infty kx_k}{2} \log \frac{x_0 + \sum_{k=1}^\infty kx_k}{2}.$$

Then

$$\int_{t_1}^{t_1+\varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds = \tilde{H}(z) + \tilde{H}(\mathbf{x}^{(2)}) - \tilde{H}(\mathbf{x}^{(1)}) + \tilde{K}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) < \infty.$$

Moreover, the right-hand side is lower semicontinuous in $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \Xi$, namely for $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \in \Xi$ with $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \rightarrow (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \Xi$,

$$\begin{aligned} \liminf_{n \rightarrow \infty} (\tilde{H}(z^n) + \tilde{H}(\mathbf{x}^{(2),n}) - \tilde{H}(\mathbf{x}^{(1),n}) + \tilde{K}(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})) \\ \geq \tilde{H}(z) + \tilde{H}(\mathbf{x}^{(2)}) - \tilde{H}(\mathbf{x}^{(1)}) + \tilde{K}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}), \end{aligned}$$

where $z^n = \mathbf{x}^{(1),n} - \mathbf{x}^{(2),n}$, $z = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$.

Recall the functions H and K from (1.3) and (2.2) respectively. We note that with \tilde{K} and \tilde{H} as introduced in the above lemma, for $\mathbf{0} \leq \mathbf{q} \leq \bar{\mathbf{q}} \leq \mathbf{p}$

$$(7.17) \quad H(\mathbf{q}) = \tilde{H}(0, \mathbf{q}), \quad K(\mathbf{q}) = \tilde{K}(0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q}).$$

The next lemma shows that $(\tilde{\xi}, \tilde{\psi})$ in Construction 1 is a minimizer for the right side of (7.12).

LEMMA 7.6. Suppose that $\varsigma \geq 0$. Suppose that $\sum_{k=1}^\infty kz_k + z_0 > 2 \sum_{k=1}^\infty z_k$. Fix $t_1 \geq 0$ and let (ξ^*, ψ^*) be as in Lemma 7.1 and $(\tilde{\xi}, \tilde{\psi})$ as introduced in Construction 1. Then

$$(7.18) \quad \begin{aligned} I_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) &= \inf_{(\xi, \psi) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})} \int_{t_1}^{t_1+\varsigma} L(\xi(s), \xi'(s)) ds \\ &= \int_{t_1}^{t_1+\varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds. \end{aligned}$$

Proofs of Lemmas 7.1–7.6 are given in Section 9. The following proposition summarizes an important consequence of the above lemmas for the case when $x_0^{(1)} = x_0^{(2)} = 0$.

PROPOSITION 7.7.

(a) Suppose $\mathbf{0} \leq \mathbf{q} \leq \bar{\mathbf{q}} \leq \mathbf{p}$ and that either $\sum_{k=1}^\infty kq_k > 2 \sum_{k=1}^\infty q_k$ or $\sum_{k=1}^\infty kq_k = 2 \sum_{k=1}^\infty q_k$ but $p_1 = 0$. Given $t_1 \geq 0$, and with $\mathbf{x}^{(0)} \doteq (0, \mathbf{p})$, $\mathbf{x}^{(1)} \doteq (0, \bar{\mathbf{q}})$, suppose there exists some $(\xi^*, \psi^*) \in \mathcal{J}_{0, t_1}^0(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})$ such that $I_{0, t_1}(\xi^*, \psi^*) < \infty$. Then

$$\begin{aligned} \inf_{t_2 \geq t_1} I_{t_1, t_2}^1((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})) &= I_{t_1, t_1+\tau}^2((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})) \\ &= H(\mathbf{q}) + H(\bar{\mathbf{q}} - \mathbf{q}) - H(\bar{\mathbf{q}}) + K(\mathbf{q}), \end{aligned}$$

where $\tau \doteq \varsigma((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})) = \frac{1}{2} \sum_{k=1}^\infty kq_k$.

(b) Suppose $p_1 = 0$, $\mathbf{q} \geq 0$, $\bar{\mathbf{q}} \geq \mathbf{0}$, $\mathbf{q} + \bar{\mathbf{q}} \leq \mathbf{p}$, $\sum_{k=1}^\infty kq_k \geq 2 \sum_{k=1}^\infty q_k$, and $\sum_{k=1}^\infty k\bar{q}_k \geq 2 \sum_{k=1}^\infty \bar{q}_k$. Let $\tau \doteq \varsigma((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})) = \frac{1}{2} \sum_{k=1}^\infty kq_k$ and $\bar{\tau} \doteq \frac{1}{2} \sum_{k=1}^\infty k\bar{q}_k$. Then

$$\begin{aligned} I_{0, \bar{\tau}}^2((0, \mathbf{p}), (0, \mathbf{p} - \bar{\mathbf{q}})) + I_{\bar{\tau}, \bar{\tau}+\tau}^2((0, \mathbf{p} - \bar{\mathbf{q}}), (0, \mathbf{p} - \bar{\mathbf{q}} - \mathbf{q})) \\ = I_{0, \tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})) + I_{\tau, \tau+\bar{\tau}}^2((0, \mathbf{p} - \mathbf{q}), (0, \mathbf{p} - \mathbf{q} - \bar{\mathbf{q}})). \end{aligned}$$

PROOF. (a) The first equality in part (a) is a consequence of Lemma 7.1. For the second equality, consider first the case $\sum_{k=1}^{\infty} kq_k > 2 \sum_{k=1}^{\infty} q_k$. From (7.17) we have

$$\begin{aligned} & H(\mathbf{q}) + H(\bar{\mathbf{q}} - \mathbf{q}) - H(\bar{\mathbf{q}}) + K(\mathbf{q}) \\ &= \tilde{H}(0, \mathbf{q}) + \tilde{H}(0, \bar{\mathbf{q}} - \mathbf{q}) - \tilde{H}(0, \bar{\mathbf{q}}) + \tilde{K}((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})). \end{aligned}$$

Applying Lemma 7.5 with $\mathbf{x}^{(1)} = (0, \bar{\mathbf{q}})$, $\mathbf{x}^{(2)} = (0, \bar{\mathbf{q}} - \mathbf{q})$, the above expression equals $\int_{t_1}^{t_1+\tau} L(\tilde{\boldsymbol{\xi}}(s), \tilde{\boldsymbol{\xi}}'(s)) ds$ where $\tilde{\boldsymbol{\xi}}$ is defined by (7.14)–(7.16). Now from Lemma 7.6

$$I_{t_1, t_1+\tau}^2((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})) = H(\mathbf{q}) + H(\bar{\mathbf{q}} - \mathbf{q}) - H(\bar{\mathbf{q}}) + K(\mathbf{q})$$

which proves the second equality in part (a) for the considered case.

Now we consider the case $\sum_{k=1}^{\infty} kq_k = 2 \sum_{k=1}^{\infty} q_k$ and $p_1 = 0$. Since $p_1 = 0$, we must have $q_k = 0$ for each $k \neq 2$. Then for any $(\boldsymbol{\xi}, \psi) \in \mathcal{J}_{t_1, t_1+\tau}^2((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q}))$ with $I_{t_1, t_1+\tau}(\boldsymbol{\xi}, \psi) < \infty$, we must have (see (2.15) and the definition of \mathcal{J}_{t_1, t_2}^2) $\zeta_2'(t) = -1$ and $\zeta_k'(t) = \psi'(t) = 0$, $k \neq 2$ for $t \in [t_1, t_1 + \tau]$. Also, in this case $q_1 = 0$ and so we are in case (i) of Construction 1 with $\mathbf{x}^{(1)} = (0, \bar{\mathbf{q}})$ and $\mathbf{x}^{(2)} = (0, \bar{\mathbf{q}} - \mathbf{q})$. It is easily checked that any $(\boldsymbol{\xi}, \psi)$ with the above properties is the same as the minimizer $(\tilde{\boldsymbol{\xi}}, \tilde{\psi})$ over $[t_1, t_1 + \tau]$. Thus using Lemma 7.1 and Lemma 7.5 we get

$$\begin{aligned} I_{t_1, t_1+\tau}^2((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})) &= \inf_{(\boldsymbol{\xi}, \psi) \in \mathcal{J}_{t_1, t_1+\tau}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})} \int_{t_1}^{t_1+\tau} L(\boldsymbol{\xi}(s), \boldsymbol{\xi}'(s)) ds \\ &= \int_{t_1}^{t_1+\tau} L(\tilde{\boldsymbol{\xi}}(s), \tilde{\boldsymbol{\xi}}'(s)) ds \\ &= \tilde{H}(0, \mathbf{q}) + \tilde{H}(0, \bar{\mathbf{q}} - \mathbf{q}) - \tilde{H}(0, \bar{\mathbf{q}}) + \tilde{K}((0, \bar{\mathbf{q}}), (0, \bar{\mathbf{q}} - \mathbf{q})) \\ &= H(\mathbf{q}) + H(\bar{\mathbf{q}} - \mathbf{q}) - H(\bar{\mathbf{q}}) + K(\mathbf{q}). \end{aligned}$$

This proves part (a) in this case.

(b) From part (a),

$$I_{0, \bar{\tau}}^2((0, \mathbf{p}), (0, \mathbf{p} - \bar{\mathbf{q}})) = H(\bar{\mathbf{q}}) + H(\mathbf{p} - \bar{\mathbf{q}}) - H(\mathbf{p}) + K(\bar{\mathbf{q}})$$

and since the right side is finite, again from part (a),

$$I_{\bar{\tau}, \bar{\tau}+\tau}^2((0, \mathbf{p} - \bar{\mathbf{q}}), (0, \mathbf{p} - \bar{\mathbf{q}} - \mathbf{q})) = H(\mathbf{q}) + H(\mathbf{p} - \bar{\mathbf{q}} - \mathbf{q}) - H(\mathbf{p} - \bar{\mathbf{q}}) + K(\mathbf{q}).$$

Therefore,

$$\begin{aligned} & I_{0, \bar{\tau}}^2((0, \mathbf{p}), (0, \mathbf{p} - \bar{\mathbf{q}})) + I_{\bar{\tau}, \bar{\tau}+\tau}^2((0, \mathbf{p} - \bar{\mathbf{q}}), (0, \mathbf{p} - \bar{\mathbf{q}} - \mathbf{q})) \\ (7.19) \quad &= [H(\bar{\mathbf{q}}) + H(\mathbf{p} - \bar{\mathbf{q}}) - H(\mathbf{p}) + K(\bar{\mathbf{q}})] \\ &+ [H(\mathbf{q}) + H(\mathbf{p} - \bar{\mathbf{q}} - \mathbf{q}) - H(\mathbf{p} - \bar{\mathbf{q}}) + K(\mathbf{q})] \\ &= [H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}) + K(\mathbf{q})] \\ &+ [H(\bar{\mathbf{q}}) + H(\mathbf{p} - \mathbf{q} - \bar{\mathbf{q}}) - H(\mathbf{p} - \mathbf{q}) + K(\bar{\mathbf{q}})] \\ &= I_{0, \bar{\tau}}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})) + I_{\bar{\tau}, \bar{\tau}+\bar{\tau}}^2((0, \mathbf{p} - \mathbf{q}), (0, \mathbf{p} - \mathbf{q} - \bar{\mathbf{q}})), \end{aligned}$$

where the last line follows, once more, from (a). This proves (b) and completes the proof. \square

8. Proof of Theorem 2.6. In this section we will use Theorem 2.16 and the results in Section 7 to prove Theorem 2.6. Let $0 \leq \mathbf{q} \leq \mathbf{p}$. Recall the (distributionally equivalent) representation of the event $E^{n,\varepsilon}(\mathbf{q})$ given in (7.1), in terms of X^n . Define

$$\begin{aligned}
 (8.1) \quad E^{n,\varepsilon,T}(\mathbf{q}) &\doteq \left\{ \exists t_1, t_2 \in [0, T] \text{ such that } X_0^n(t_1-) = X_0^n(t_2) = -1/n, \right. \\
 &\quad X_0^n(t) > -1/n \text{ for } t \in [t_1, t_2], \\
 &\quad |X_k^n(t_1-) - X_k^n(t_2) - q_k| \leq \varepsilon \text{ for all } k \in \mathbb{N} \} \\
 &= \left\{ \exists t_1, t_2 \in [0, T] \text{ such that } X_0^n(t_1-) = X_0^n(t_2) = -1/n, \right. \\
 &\quad Y^n(t) > Y^n(t_1-) - 2/n \text{ for } t \in [t_1, t_2], \\
 &\quad \left. |X_k^n(t_1-) - X_k^n(t_2) - q_k| \leq \varepsilon \text{ for all } k \in \mathbb{N} \right\}.
 \end{aligned}$$

Note that $E^{n,\varepsilon,T}(\mathbf{q}) \subset E^{n,\varepsilon}(\mathbf{q})$ but they are not equal, since the continuous-time EEA may not terminate by time T . Consider the event that the continuous-time EEA terminates before time T , namely the event $F^{n,T}$ defined as

$$(8.2) \quad F^{n,T} \doteq \{X^n(T) = (-1/n, \mathbf{0})\}.$$

Then

$$(8.3) \quad E^{n,\varepsilon}(\mathbf{q}) \cap F^{n,T} \subset E^{n,\varepsilon,T}(\mathbf{q}) \subset E^{n,\varepsilon}(\mathbf{q}).$$

The following lemma guarantees that in order to study the exponential rate of decay of $P(E^{n,\varepsilon}(\mathbf{q}))$, it suffices to study that of $P(E^{n,\varepsilon,T}(\mathbf{q}))$.

LEMMA 8.1. $\limsup_{n \rightarrow \infty} \frac{1}{n} \log P((F^{n,T})^c) \rightarrow -\infty$ as $T \rightarrow \infty$.

PROOF. Recall from Section 2.3 that the discrete-time EEA terminates in at most nN steps where $N \doteq \lfloor \sup_n \frac{1}{2} \sum_{k=1}^{\infty} k \frac{n_k}{n} \rfloor + 1 < \infty$. So since the discrete time EEA is the embedded chain associated with the continuous time EEA (see Section 2.4), X^n will have at most nN jumps before arriving at the absorbing state $(-\frac{1}{n}, \mathbf{0})$. Since the total jump rate for $X^n(t)$ at any instant before getting absorbed is $n \sum_{k=0}^{\infty} r_k(X^n(t)) = n$, we have

$$P(F^{n,T}) \geq P\left(\sum_{i=1}^{nN} \xi_i \leq T\right) = P\left(\frac{1}{n} \sum_{i=1}^{nN} \tilde{\xi}_i \leq T\right),$$

where ξ_i are i.i.d. $\exp(n)$ and $\tilde{\xi}_i$ are i.i.d. $\exp(1)$. Therefore,

$$\begin{aligned}
 \limsup_{n \rightarrow \infty} \frac{1}{n} \log P((F^{n,T})^c) &\leq N \limsup_{n \rightarrow \infty} \frac{1}{nN} \log \mathbb{P}\left(\frac{1}{nN} \sum_{i=1}^{nN} \tilde{\xi}_i > \frac{T}{N}\right) \\
 &= -NL_1\left(\frac{T}{N}\right) \rightarrow -\infty
 \end{aligned}$$

as $T \rightarrow \infty$, where the second equality is from Cramér's theorem and where $L_1(x) \doteq x - 1 - \log x$ for $x \geq 0$ is the Legendre transform of the log-moment generating function of $\tilde{\xi}_1$. \square

The following lemma gives an upper bound for the exponential rate of decay of $P(E^{n,\varepsilon}(\mathbf{q}))$.

LEMMA 8.2. Suppose $\sum_{k=1}^{\infty} kq_k > 2 \sum_{k=1}^{\infty} q_k$. Then the upper bound in (7.5) holds, namely

$$(8.4) \quad \begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \\ & \leq - \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0,t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1,t_1+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))], \end{aligned}$$

where $\tau \doteq \varsigma((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q})) = \frac{1}{2} \sum_{k=1}^{\infty} kq_k$ for each $\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}$.

PROOF. From (8.3) we have

$$P(E^{n,\varepsilon}(\mathbf{q})) \leq P(E^{n,\varepsilon,T}(\mathbf{q})) + P((F^{n,T})^c)$$

and hence

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \leq \max \left\{ \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon,T}(\mathbf{q})), \limsup_{n \rightarrow \infty} \frac{1}{n} \log P((F^{n,T})^c) \right\}.$$

In view of Lemma 8.1, it suffices to show that for all sufficiently large T

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon,T}(\mathbf{q})) \\ & \leq - \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0,t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1,t_1+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))]. \end{aligned}$$

Let $\mathfrak{P}_T \doteq \mathbb{D}([0, T] : \mathbb{R} \times \mathbb{R}_+^\infty \times \mathbb{R})$ and consider

$$\begin{aligned} \tilde{E}^{\varepsilon,T}(\mathbf{q}) & \doteq \{(\boldsymbol{\zeta}, \psi) \in \mathfrak{P}_T : \exists t_1, t_2 \in [0, T] \text{ such that } \zeta_0(t_1-) = \zeta_0(t_2) \leq 0, \\ & \quad \psi(t) \geq \psi(t_1-) - \varepsilon \text{ for } t \in [t_1, t_2], |\zeta_k(t_1-) - \zeta_k(t_2) - q_k| \leq \varepsilon \text{ for all } k \in \mathbb{N}\}. \end{aligned}$$

Denote the closure of $\tilde{E}^{\varepsilon,T}(\mathbf{q})$ by $cl\tilde{E}^{\varepsilon,T}(\mathbf{q})$. From the definition in (8.1), when $n > 2\varepsilon^{-1}$

$$E^{n,\varepsilon,T}(\mathbf{q}) \subset \{(X^n, Y^n) \in \tilde{E}^{\varepsilon,T}(\mathbf{q})\} \subset \{(X^n, Y^n) \in cl\tilde{E}^{\varepsilon,T}(\mathbf{q})\}.$$

From this and Theorem 2.16 we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon,T}(\mathbf{q})) & \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \log P((X^n, Y^n) \in cl\tilde{E}^{\varepsilon,T}(\mathbf{q})) \\ & \leq - \inf_{(\boldsymbol{\zeta}, \psi) \in cl\tilde{E}^{\varepsilon,T}(\mathbf{q})} I_T(\boldsymbol{\zeta}, \psi). \end{aligned}$$

Since $I_T(\boldsymbol{\zeta}, \psi) < \infty$ only when $(\boldsymbol{\zeta}, \psi) \in \mathcal{C}_T$, we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon,T}(\mathbf{q})) \leq - \inf_{(\boldsymbol{\zeta}, \psi) \in cl\tilde{E}^{\varepsilon,T}(\mathbf{q}) \cap \mathcal{C}_T} I_T(\boldsymbol{\zeta}, \psi).$$

It is easy to see that $cl\tilde{E}^{\varepsilon,T}(\mathbf{q}) \cap \mathcal{C}_T = \tilde{E}^{\varepsilon,T}(\mathbf{q}) \cap \mathcal{C}_T$. Thus we have

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon,T}(\mathbf{q})) \leq - \inf_{(\boldsymbol{\zeta}, \psi) \in \tilde{E}^{\varepsilon,T}(\mathbf{q}) \cap \mathcal{C}_T} I_T(\boldsymbol{\zeta}, \psi).$$

Letting

$$\tilde{E}^T(\mathbf{q}) \doteq \{(\boldsymbol{\zeta}, \psi) \in \mathcal{C}_T : \exists t_1, t_2 \in [0, T] \text{ such that } \zeta_0(t_1) = \zeta_0(t_2) \leq 0,$$

$$\psi(t) \geq \psi(t_1) \text{ for } t \in [t_1, t_2], \zeta_k(t_1) - \zeta_k(t_2) = q_k \text{ for all } k \in \mathbb{N}\},$$

we have $\tilde{E}^T(\mathbf{q}) = \bigcap_{\varepsilon > 0} (\tilde{E}^{\varepsilon, T}(\mathbf{q}) \cap \mathcal{C}_T)$. From this, the lower semicontinuity and compactness of level sets of $I_T(\boldsymbol{\zeta}, \psi)$ (since I_T is a rate function; see Theorem 2.16), it follows that

$$\begin{aligned} \limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n, \varepsilon, T}(\mathbf{q})) &\leq -\liminf_{\varepsilon \rightarrow 0} \inf_{(\boldsymbol{\zeta}, \psi) \in \tilde{E}^{\varepsilon, T}(\mathbf{q}) \cap \mathcal{C}_T} I_T(\boldsymbol{\zeta}, \psi) \\ &= -\inf_{(\boldsymbol{\zeta}, \psi) \in \tilde{E}^T(\mathbf{q})} I_T(\boldsymbol{\zeta}, \psi). \end{aligned}$$

Breaking $(\boldsymbol{\zeta}, \psi) \in \tilde{E}^T(\mathbf{q})$ for $t \in [0, T]$ according to $t \leq t_1$, $t_1 \leq t \leq t_2$ and $t \geq t_2$,

$$\begin{aligned} \inf_{(\boldsymbol{\zeta}, \psi) \in \tilde{E}^T(\mathbf{q})} I_T(\boldsymbol{\zeta}, \psi) &= \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, 0 \leq t_1 < t_2 \leq T} [I_{0, t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1, t_2}^1((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))] \\ &\geq \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, 0 \leq t_1 < t_2 < \infty} [I_{0, t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1, t_2}^1((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))] \\ &= \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0, t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1, t_1 + \tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))], \end{aligned}$$

where the last line follows from Lemma 7.1. The result follows. \square

The following lemma improves the upper bound (8.4) in Lemma 8.2 when $p_1 = 0$.

LEMMA 8.3. *Suppose $p_1 = 0$ and $\mathbf{0} \leq \mathbf{q} \leq \mathbf{p}$. Then:*

- (a) $I_{0, t_1}^0((0, \mathbf{p}), (0, \mathbf{q})) = I_{0, t_1}^1((0, \mathbf{p}), (0, \mathbf{q}))$ for each $t_1 \geq 0$.
- (b) Let $\tau \doteq \frac{1}{2} \sum_{k=1}^{\infty} k q_k$ as in Lemma 8.2. The infimum on the right side of (8.4) is achieved at $t_1 = 0$:

$$\inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0, t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1, t_1 + \tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))] = I_{0, \tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})).$$

PROOF. (a) Fix $t_1 \geq 0$. It suffices to show that if $(\boldsymbol{\zeta}, \psi) \in \mathcal{J}_{0, t_1}^0((0, \mathbf{p}), (0, \mathbf{q}))$ satisfies $I_{0, t_1}(\boldsymbol{\zeta}, \psi) < \infty$ then $(\boldsymbol{\zeta}, \psi) \in \mathcal{J}_{0, t_1}^1((0, \mathbf{p}), (0, \mathbf{q}))$. For such a pair of $(\boldsymbol{\zeta}, \psi)$, let $\varphi \in \mathcal{S}_{t_1}(\boldsymbol{\zeta}, \psi)$ be such that the associated cost is finite. In particular ψ , and consequently ζ_0 , is absolutely continuous.

Since $\zeta_0(t) = \Gamma(\psi)(t) \geq 0$ for $t \in [0, t_1]$, we have

$$(8.5) \quad 1_{\{\zeta_0(t) > 0\}} \zeta_0'(t) = 1_{\{\zeta_0(t) > 0\}} \psi'(t), \quad 1_{\{\zeta_0(t) = 0\}} \zeta_0'(t) = 0, \quad \text{a.e. } t \in [0, t_1].$$

Since $p_1 = 0$, we see from (2.15) that $1_{\{\zeta_0(t) = 0\}} \psi'(t) \geq 0$ for a.e. $t \in [0, t_1]$. Indeed, when $p_1 = 0$ the term for $k = 1$ in the sum on the right side of (2.15) is zero. Also, the term for $k = 2$ is always zero and the integrand for $k = 0$ is zero on the set $\{\zeta_0(t) = 0\}$. This shows that, on this set, the derivative of the sum on the right side of (2.15) is nonnegative. Combining this with (8.5), we have for $t \in [0, t_1]$,

$$\begin{aligned} \psi(t) &= \int_0^t 1_{\{\zeta_0(s) > 0\}} \psi'(s) ds + \int_0^t 1_{\{\zeta_0(s) = 0\}} \psi'(s) ds \\ &\geq \int_0^t 1_{\{\zeta_0(s) > 0\}} \zeta_0'(s) ds \\ &= \int_0^t 1_{\{\zeta_0(s) > 0\}} \zeta_0'(s) ds + \int_0^t 1_{\{\zeta_0(s) = 0\}} \zeta_0'(s) ds = \zeta_0(t) \geq 0. \end{aligned}$$

This implies $(\boldsymbol{\zeta}, \psi) \in \mathcal{J}_{0, t_1}^1((0, \mathbf{p}), (0, \mathbf{q}))$ and part (a) follows.

(b) For $\mathbf{0} \leq \bar{\mathbf{p}} \leq \mathbf{p}$, let $\bar{\mathbf{q}} \doteq \mathbf{p} - \bar{\mathbf{p}}$ and $\bar{\tau} \doteq \frac{1}{2} \sum_{k=1}^{\infty} k \bar{q}_k$. Since $p_1 = 0$, we always have $\sum_{k=1}^{\infty} k q_k \geq 2 \sum_{k=1}^{\infty} q_k$ and $\sum_{k=1}^{\infty} k \bar{q}_k \geq 2 \sum_{k=1}^{\infty} \bar{q}_k$. Therefore,

$$\begin{aligned} & \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0,t_1}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1,t_1+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))] \\ &= \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}, t_1 \geq 0} [I_{0,t_1}^1((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{t_1,t_1+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))] \\ &= \inf_{\mathbf{q} \leq \bar{\mathbf{p}} \leq \mathbf{p}} [I_{0,\bar{\tau}}^2((0, \mathbf{p}), (0, \bar{\mathbf{p}})) + I_{\bar{\tau},\bar{\tau}+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))] \\ &= \inf_{\mathbf{0} \leq \bar{\mathbf{q}} \leq \mathbf{p} - \mathbf{q}} [I_{0,\bar{\tau}}^2((0, \mathbf{p}), (0, \mathbf{p} - \bar{\mathbf{q}})) + I_{\bar{\tau},\bar{\tau}+\tau}^2((0, \mathbf{p} - \bar{\mathbf{q}}), (0, \mathbf{p} - \bar{\mathbf{q}} - \mathbf{q}))] \end{aligned}$$

where the first equality uses part (a) with $\mathbf{q} = \bar{\mathbf{p}}$ and the second equality follows from Lemma 7.1 and the observation that $I_{t,t+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q})) = I_{t',t'+\tau}^2((0, \bar{\mathbf{p}}), (0, \bar{\mathbf{p}} - \mathbf{q}))$ for all t, t' as long as $I_{0,s}^0((0, \mathbf{p}), (0, \bar{\mathbf{p}})) < \infty$ for $s = t, t'$. Using Proposition 7.7(b), the right side on the last line equals

$$\begin{aligned} & \inf_{\mathbf{0} \leq \bar{\mathbf{q}} \leq \mathbf{p} - \mathbf{q}} [I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})) + I_{\tau,\tau+\bar{\tau}}^2((0, \mathbf{p} - \mathbf{q}), (0, \mathbf{p} - \mathbf{q} - \bar{\mathbf{q}}))] \\ &= I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})) + \inf_{\mathbf{0} \leq \bar{\mathbf{q}} \leq \mathbf{p} - \mathbf{q}} I_{\tau,\tau+\bar{\tau}}^2((0, \mathbf{p} - \mathbf{q}), (0, \mathbf{p} - \mathbf{q} - \bar{\mathbf{q}})) \\ &= I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})), \end{aligned}$$

where the last equality follows by considering $\bar{\mathbf{q}} = 0$. This completes the proof. \square

Next we will prove the lower bound.

LEMMA 8.4. Suppose $\mathbf{0} \leq \mathbf{q} \leq \mathbf{p}$ and $\sum_{k=1}^{\infty} k q_k > 2 \sum_{k=1}^{\infty} q_k$. Let $\tau \doteq \frac{1}{2} \sum_{k=1}^{\infty} k q_k$. Then the lower bound in (7.4) holds.

PROOF. Let $(\tilde{\xi}(t), \tilde{\psi}(t))$ be as introduced in Construction 1 for $t \leq \tau$, with $t_1 = 0$, $\mathbf{x}^{(1)} = (0, \mathbf{p})$ and $\mathbf{x}^{(2)} = (0, \mathbf{p} - \mathbf{q})$. We define $(\tilde{\xi}(t), \tilde{\psi}(t))$ for $t > \tau$ through (2.15)–(2.16) by setting $\varphi_k(t, y) = 1$ for all k, y and $t > \tau$. Then $I_t(\tilde{\xi}, \tilde{\psi}) = I_{\tau}(\tilde{\xi}, \tilde{\psi})$ for all $t > \tau$. So by Lemmas 7.1 and 7.6, for $t \geq \tau$

$$(8.6) \quad I_t(\tilde{\xi}, \tilde{\psi}) = I_{\tau}(\tilde{\xi}, \tilde{\psi}) = \int_0^{\tau} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds = I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})).$$

For $\delta \in (0, 1)$ consider the set

$$\begin{aligned} (8.7) \quad \tilde{G}_{\delta}(\tilde{\xi}, \tilde{\psi}) & \doteq \left\{ (\xi, \psi) \in \mathbb{D}([0, \infty) : \mathbb{R} \times \mathbb{R}_+^{\infty} \times \mathbb{R}) : \sup_{t \in [0, \tau]} |\xi_k(t) - \tilde{\xi}_k(t)| < \delta, \right. \\ & \quad \left. \text{for all } k = 0, 1, 2, \dots, \lfloor \delta^{-1} \rfloor \right\}. \end{aligned}$$

Let $\tau^n \doteq \inf\{t \geq \tau : X_0^n(t) = -\frac{1}{n}\}$. Then $\tau^n < \infty$ a.s. Define for odd integer $j \geq -1$,

$$G_j^n \doteq \left\{ X_0^n(\tau) = \frac{j}{n}, X_k^n(\tau^n) = X_k^n(\tau), k \in \mathbb{N} \right\},$$

and for even integer $j \geq -1$,

$$G_j^n \doteq \left\{ X_0^n(\tau) = \frac{j}{n}, \sum_{k=1}^{\infty} (X_k^n(\tau^n) - X_k^n(\tau)) = -\frac{1}{n} \right\}.$$

Intuitively, G_j^n describes the event that from time instant τ to the time τ^n at which the current component is fully explored, the continuous-time EEA does not wake up any sleeping vertices, with the exception that if the number of active half-edges at time τ is odd (namely $X_0^n(\tau) = \frac{j}{n}$ for some even integer $j \geq -1$), in which case exactly one sleeping vertex (necessarily with odd degree) will be woken up. Consider the event

$$A_\delta^n(\tilde{\xi}, \tilde{\psi}) \doteq \{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\} \cap \left(\bigcup_{j=-1}^{\infty} G_j^n \right).$$

Fix $\varepsilon \in (0, 1)$. We claim that there exist $\delta_0 > 0$ and $n_0 > 0$ such that

$$(8.8) \quad A_\delta^n(\tilde{\xi}, \tilde{\psi}) \subset E^{n, \varepsilon}(\mathbf{q}) \quad \text{for all } \delta < \delta_0 \text{ and } n > n_0.$$

To see this, first note that by Assumption 2.2, there exists $M \in \mathbb{N}$ such that

$$(8.9) \quad \sup_{n \in \mathbb{N}} \sum_{k=M}^{\infty} k \frac{n_k}{n} < \frac{\varepsilon}{2}, \quad \sum_{k=M}^{\infty} kp_k < \frac{\varepsilon}{2}.$$

By continuity of $\tilde{\xi}$, there exists $\varepsilon_0 > 0$ such that

$$(8.10) \quad |\tilde{\xi}_k(t) - \tilde{\xi}_k(0)| < \frac{\varepsilon}{4} \quad \text{for all } t \in [0, \varepsilon_0], k = 0, 1, \dots, M,$$

$$(8.11) \quad |\tilde{\xi}_k(t) - \tilde{\xi}_k(\tau)| < \frac{\varepsilon}{4} \quad \text{for all } t \in [\tau - \varepsilon_0, \tau], k = 0, 1, \dots, M.$$

From Lemma 7.4(c) we have $\tilde{\xi}_0(t) > 0$ for all $t \in (0, \tau)$. Since $\tilde{\xi}_0(t)$ is continuous,

$$\delta_0 \doteq \left(\inf_{t \in [\varepsilon_0, \tau - \varepsilon_0]} \tilde{\xi}_0(t) \right) \wedge \frac{\varepsilon}{4} \wedge \frac{1}{M} > 0.$$

Take $n_0 > \frac{4}{\varepsilon}$. We now show (8.8) with this choice of n_0 and δ_0 . Fix $\delta < \delta_0$ and $n > n_0$ and consider $\omega \in A_\delta^n(\tilde{\xi}, \tilde{\psi})$. For $t \in [\varepsilon_0, \tau - \varepsilon_0]$, since $|X_0^n(t) - \tilde{\xi}_0(t)| < \delta < \delta_0 \leq \tilde{\xi}_0(t)$, we have $\inf_{t \in [\varepsilon_0, \tau - \varepsilon_0]} X_0^n(t) > 0$. So there exist $t_1^n \in [0, \varepsilon_0]$ and $t_2^n \in [\tau - \varepsilon_0, \tau^n]$ such that

$$(8.12) \quad X_0^n(t_1^n) = X_0^n(t_2^n) = -\frac{1}{n}, \quad X_0^n(t) > -\frac{1}{n} \quad \text{for } t \in [t_1^n, t_2^n],$$

where by convention $X_0^n(0-) = X_0^n(0) = -1/n$. For $k \geq M$, it follows from (8.9) that

$$(8.13) \quad |X_k^n(t_1^n) - X_k^n(t_2^n) - q_k| \leq |X_k^n(t_1^n) - X_k^n(t_2^n)| + q_k \leq \frac{n_k}{n} + p_k < \varepsilon.$$

For $1 \leq k \leq M \leq \lfloor \delta^{-1} \rfloor$,

$$(8.14) \quad \begin{aligned} |X_k^n(t_1^n) - X_k^n(t_2^n) - q_k| &= |(X_k^n(t_1^n) - X_k^n(t_2^n)) - (\tilde{\xi}_k(0) - \tilde{\xi}_k(\tau))| \\ &\leq |X_k^n(t_1^n) - \tilde{\xi}_k(0)| + |X_k^n(t_2^n) - \tilde{\xi}_k(\tau)|. \end{aligned}$$

From (8.7) and (8.10) we have the following bound for the first term in (8.14):

$$|X_k^n(t_1^n) - \tilde{\xi}_k(0)| \leq |X_k^n(t_1^n) - \tilde{\xi}_k(t_1^n)| + |\tilde{\xi}_k(t_1^n) - \tilde{\xi}_k(0)| < \delta + \frac{\varepsilon}{4}.$$

For the second term in (8.14), if $t_2^n \leq \tau$, then using (8.7) and (8.11) we have

$$|X_k^n(t_2^n) - \tilde{\xi}_k(\tau)| \leq |X_k^n(t_2^n) - \tilde{\xi}_k(t_2^n)| + |\tilde{\xi}_k(t_2^n) - \tilde{\xi}_k(\tau)| < \delta + \frac{\varepsilon}{4}.$$

If $t_2^n > \tau$, then $t_2^n = \tau^n$ and from the definition of G_j^n and (8.7) we have

$$|X_k^n(t_2^n) - \tilde{\xi}_k(\tau)| \leq |X_k^n(\tau^n) - X_k^n(\tau)| + |X_k^n(\tau) - \tilde{\xi}_k(\tau)| \leq \frac{1}{n} + \delta \leq \frac{\varepsilon}{4} + \delta.$$

Combining these three displays with (8.14) gives

$$|X_k^n(t_1^n) - X_k^n(t_2^n) - q_k| < 2\left(\delta + \frac{\varepsilon}{4}\right) < \varepsilon, \quad k \in \mathbb{N}.$$

From this, and (7.1), (8.12), (8.13) we have $\omega \in E^{n,\varepsilon}(\mathbf{q})$. Since $\delta < \delta_0$ and $n > n_0$ is arbitrary, the claim (8.8) holds.

For fixed $\delta < \delta_0$ and $n > n_0$ consider the following two probabilities

$$P(A_\delta^n(\tilde{\xi}, \tilde{\psi})), \quad P((X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})).$$

Write

$$\begin{aligned} P(A_\delta^n(\tilde{\xi}, \tilde{\psi})) &= \sum_{j=-1}^{\infty} P(\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\} \cap G_j^n) \\ (8.15) \quad &= \sum_{j=-1}^{\lfloor \delta n \rfloor} \mathbf{E}[1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} P(G_j^n | \mathcal{F}_\tau)], \end{aligned}$$

where we only have to sum up to $\lfloor \delta n \rfloor$ in the last line when $(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})$ since $\tilde{\zeta}_0(\tau) = 0$. Since $\tau^n = \tau$ on $\{X_0^n(\tau) = -\frac{1}{n}\}$, we have

$$(8.16) \quad P(G_{-1}^n | \mathcal{F}_\tau) = 1_{\{X_0^n(\tau) = -\frac{1}{n}\}}.$$

From Assumption 2.2 and the fact that $r(X^n(t))$ is nonincreasing it follows

$$\sup_{n \in \mathbb{N}} \sup_{t \geq 0} r(X^n(t)) \leq \sup_{n \in \mathbb{N}} \sum_{k=1}^{\infty} \frac{kn_k}{n} \doteq C_0 < \infty.$$

Hence for odd integer $1 \leq j \leq \lfloor \delta n \rfloor$ and $\delta < \frac{C_0}{2}$,

$$\begin{aligned} &1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} P(G_j^n | \mathcal{F}_\tau) \\ &= 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{j}{nr(X^n(\tau))} \\ (8.17) \quad &\cdot \frac{j-2}{nr(X^n(\tau)) - 2} \cdots \frac{1}{nr(X^n(\tau)) - (j-1)} 1_{\{X_0^n(\tau) = \frac{j}{n}\}} \\ &\geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{j}{C_0 n} \cdot \frac{j-2}{C_0 n} \cdots \frac{1}{C_0 n} 1_{\{X_0^n(\tau) = \frac{j}{n}\}} \\ &\geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{\lfloor 2\delta n \rfloor!}{(C_0 n)^{\lfloor 2\delta n \rfloor}} 1_{\{X_0^n(\tau) = \frac{j}{n}\}}, \end{aligned}$$

where the last inequality follows since the term on the last line includes more fractions that are less than 1 than the one on the previous line. For the first equality we have used the fact that on the event G_j^n all the active $j+1$ half edges (an even number) at time instant τ should merge among themselves (without waking any sleeping vertices) by the time instant τ^n , whereas the total number of available half edges (either awake or sleeping) at time instant τ equals $nr(X^n(\tau)) + 1$.

For even integer $0 \leq j \leq \lfloor \delta n \rfloor$, we consider three different cases for values of \mathbf{p} and \mathbf{q} .

Case 1: There exists some odd $m \in \mathbb{N}$ such that $p_m > q_m \geq 0$. Let $C_m \doteq \frac{1}{2}(p_m - q_m) > 0$. For $\delta < \frac{1}{m} \wedge \delta_0 \wedge C_m$ and $(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})$, we have from (8.7)

$$X_m^n(\tau) = \tilde{\xi}_m^n(\tau) - (\tilde{\zeta}_m^n(\tau) - X_m^n(\tau)) > (p_m - q_m) - \delta > C_m,$$

which implies $X_m^n(\tau) \geq 1/n$ for $n \geq \delta^{-1}$. So for even integer $0 \leq j \leq \lfloor \delta n \rfloor$ and $n > \frac{m}{\delta} \vee n_0$,

$$\begin{aligned}
& 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} P(G_j^n | \mathcal{F}_\tau) \\
& \geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{mnX_m^n(\tau)}{nr(X^n(\tau))} \cdot \frac{j+m-2}{nr(X^n(\tau))-2} \cdot \frac{j+m-4}{nr(X^n(\tau))-4} \\
(8.18) \quad & \cdots \frac{1}{nr(X^n(\tau))-(j+m-1)} 1_{\{X_0^n(\tau)=\frac{j}{n}\}} \\
& \geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{m}{C_0n} \cdot \frac{j+m-2}{C_0n} \cdot \frac{j+m-4}{C_0n} \cdots \frac{1}{C_0n} 1_{\{X_0^n(\tau)=\frac{j}{n}\}} \\
& \geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{\lfloor 2\delta n \rfloor!}{(C_0n)^{\lfloor 2\delta n \rfloor}} 1_{\{X_0^n(\tau)=\frac{j}{n}\}},
\end{aligned}$$

where the last inequality follows once again as in (8.17). Combining (8.15)–(8.18) implies that, for $\delta < \frac{C_0}{2} \wedge \frac{1}{m} \wedge \delta_0 \wedge C_m$ and $n > \frac{m}{\delta} \vee n_0$,

$$\begin{aligned}
P(A_\delta^n(\tilde{\xi}, \tilde{\psi})) & \geq \sum_{j=-1}^{\lfloor \delta n \rfloor} \mathbf{E} \left[1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})\}} \frac{\lfloor 2\delta n \rfloor!}{(C_0n)^{\lfloor 2\delta n \rfloor}} 1_{\{X_0^n(\tau)=\frac{j}{n}\}} \right] \\
& = \frac{\lfloor 2\delta n \rfloor!}{(C_0n)^{\lfloor 2\delta n \rfloor}} P((X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})) \\
& \geq \sqrt{2\pi \lfloor 2\delta n \rfloor} \left(\frac{\lfloor 2\delta n \rfloor}{C_0en} \right)^{\lfloor 2\delta n \rfloor} P((X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})),
\end{aligned}$$

where the last line uses Stirling's approximation $n! \geq \sqrt{2\pi n} (\frac{n}{e})^n$. From this and (8.8) we have

$$\begin{aligned}
& \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \\
& \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(A_\delta^n(\tilde{\xi}, \tilde{\psi})) \\
(8.19) \quad & \geq \liminf_{n \rightarrow \infty} \left[\frac{1}{2n} \log(2\pi \lfloor 2\delta n \rfloor) + \frac{\lfloor 2\delta n \rfloor}{n} \log \left(\frac{\lfloor 2\delta n \rfloor}{C_0en} \right) \right. \\
& \quad \left. + \frac{1}{n} \log P((X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})) \right] \\
& = 2\delta \log \left(\frac{2\delta}{C_0e} \right) + \liminf_{n \rightarrow \infty} \frac{1}{n} \log P((X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})).
\end{aligned}$$

Define the open set

$$\begin{aligned}
G_{\delta,\tau}(\tilde{\xi}, \tilde{\psi}) & \doteq \left\{ (\xi, \psi) \in \mathbb{D}([0, \tau] : \mathbb{R} \times \mathbb{R}_+^\infty \times \mathbb{R}) : \sup_{t \in [0, \tau]} |\xi_k(t) - \tilde{\xi}_k(t)| < \delta, \right. \\
& \quad \left. \text{for all } k = 0, 1, 2, \dots, \lfloor \delta^{-1} \rfloor \right\}.
\end{aligned}$$

It follows from Theorem 2.16 that

$$\begin{aligned}
\liminf_{n \rightarrow \infty} \frac{1}{n} \log P((X^n, Y^n) \in \tilde{G}_\delta(\tilde{\xi}, \tilde{\psi})) & = \liminf_{n \rightarrow \infty} \frac{1}{n} \log P((X^n, Y^n) \in G_{\delta,\tau}(\tilde{\xi}, \tilde{\psi})) \\
& \geq - \inf_{(\xi, \psi) \in G_{\delta,\tau}(\tilde{\xi}, \tilde{\psi})} I_\tau((\xi, \psi)) \geq -I_\tau(\tilde{\xi}, \tilde{\psi}).
\end{aligned}$$

Combining this with (8.19), (8.6) and sending $\delta \rightarrow 0$ gives

$$(8.20) \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \geq -I_\tau(\tilde{\boldsymbol{\xi}}, \tilde{\psi}) = -I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})).$$

The lower bound in Case 1 now follows on sending $\varepsilon \rightarrow 0$.

Case 2: $p_m = 0$ for all odd $m \in \mathbb{N}$. It suffices to establish a similar estimate as in (8.18), the lower bound in Case 2 will then follow as in Case 1. From Assumptions 2.1 and 2.2

$$\sum_{k=0}^{\infty} (2k+1) \frac{n_{2k+1}}{n} \rightarrow \sum_{k=0}^{\infty} (2k+1) p_{2k+1} = 0.$$

Therefore, for each $\kappa \in (0, 1)$, there exists some $\tilde{n}_\kappa \in \mathbb{N}$ such that $0 \leq \sum_{k=0}^{\infty} (2k+1) \frac{n_{2k+1}}{n} < \kappa$ for $n > \tilde{n}_\kappa$, which implies $n_m = 0$ for all odd $m \geq \kappa n$. Consider now an even integer $0 \leq j \leq \lfloor \delta n \rfloor$ and $n > \tilde{n}_\kappa$. Denote by M^n the largest odd degree for which there is a sleeping vertex at time instant τ in the continuous time EEA. Note that $M^n \leq \kappa n$ a.s. Therefore, for $\kappa < \delta < \frac{C_0}{2} \wedge \delta_0$ and $n > n_0 \vee \tilde{n}_\kappa$,

$$\begin{aligned} & 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\boldsymbol{\xi}}, \tilde{\psi})\}} P(G_j^n | \mathcal{F}_\tau) \\ & \geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\boldsymbol{\xi}}, \tilde{\psi})\}} \sum_{1 \leq m \leq \lfloor \kappa n \rfloor, m \text{ is odd}} \left[1_{\{M^n = m\}} \frac{mnX_m^n(\tau)}{nr(X^n(\tau))} \cdot \frac{j+m-2}{nr(X^n(\tau)) - 2} \right. \\ & \quad \cdot \frac{j+m-4}{nr(X^n(\tau)) - 4} \cdots \left. \frac{1}{nr(X^n(\tau)) - (j+m-1)} \right] 1_{\{X_0^n(\tau) = \frac{j}{n}\}}. \end{aligned}$$

Since $nX_m^n(\tau) \geq 1$ on the set $\{M^n = m\}$ the right side can be bounded below by

$$\begin{aligned} & 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\boldsymbol{\xi}}, \tilde{\psi})\}} \sum_{1 \leq m \leq \lfloor \kappa n \rfloor, m \text{ is odd}} \left[1_{\{M^n = m\}} \frac{(j+m-2)!!m}{(C_0n)^{(j+m+1)/2}} \right] 1_{\{X_0^n(\tau) = \frac{j}{n}\}} \\ & \geq 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\boldsymbol{\xi}}, \tilde{\psi})\}} \sum_{1 \leq m \leq \lfloor \kappa n \rfloor, m \text{ is odd}} \left[1_{\{M^n = m\}} \frac{\lfloor 2\delta n \rfloor!}{(C_0n)^{\lfloor 2\delta n \rfloor}} \right] 1_{\{X_0^n(\tau) = \frac{j}{n}\}} \\ & = 1_{\{(X^n, Y^n) \in \tilde{G}_\delta(\tilde{\boldsymbol{\xi}}, \tilde{\psi})\}} \frac{\lfloor 2\delta n \rfloor!}{(C_0n)^{\lfloor 2\delta n \rfloor}} 1_{\{X_0^n(\tau) = \frac{j}{n}\}}, \end{aligned}$$

where the last inequality follows once again as in (8.17). Therefore, we have the same inequality as in (8.18) for $n > n_0 \vee \tilde{n}_\kappa$ and $\delta < \frac{C_0}{2} \wedge \delta_0$, and so the lower bound in Case 2 follows.

Case 3: There exists an odd $m \in \mathbb{N}$ such that $p_m > 0$ but $p_m = q_m$. For $i \in \mathbb{N}$, consider the vector $\mathbf{q}^i \doteq (q_k^i)_{k \in \mathbb{N}}$, where $q_k^i \doteq q_k$ for $k \neq m$ and $q_m^i \doteq q_m - \frac{1}{i}$. Fix $\varepsilon \in (0, 1)$. Choose i so that $p_m > q_m^i > 0$, $\varepsilon > \frac{1}{i} = q_m - q_m^i$ and $\sum_{k=1}^{\infty} kq_k^i > 2 \sum_{k=1}^{\infty} q_k^i$. When \mathbf{q} is replaced by \mathbf{q}^i we are in Case 1 and thus for $\varepsilon^i < (\varepsilon - \frac{1}{i})$, from the lower bound (8.20) for Case 1,

$$(8.21) \quad \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon}(\mathbf{q})) \geq \liminf_{n \rightarrow \infty} \frac{1}{n} \log P(E^{n,\varepsilon^i}(\mathbf{q}^i)) \geq -I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q}^i)).$$

From Proposition 7.7(a) we have

$$(8.22) \quad I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q}^i)) = H(\mathbf{q}^i) + H(\mathbf{p} - \mathbf{q}^i) - H(\mathbf{p}) + K(\mathbf{q}^i).$$

Since $((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q}^i))$ and $((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q}))$ are in Ξ and $\mathbf{q}^i \rightarrow \mathbf{q}$, from Lemma 7.3 we have $K(\mathbf{q}^i) \rightarrow K(\mathbf{q})$. Also, clearly

$$H(\mathbf{q}^i) + H(\mathbf{p} - \mathbf{q}^i) - H(\mathbf{p}) \rightarrow H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}).$$

Thus, as $i \rightarrow \infty$, the right side of (8.22) converges to

$$H(\mathbf{q}) + H(\mathbf{p} - \mathbf{q}) - H(\mathbf{p}) + K(\mathbf{q}) = I_{0,\tau}^2((0, \mathbf{p}), (0, \mathbf{p} - \mathbf{q})),$$

where the equality follows again by Proposition 7.7(a). The desired result now follows on sending $i \rightarrow \infty$ and then $\varepsilon \rightarrow 0$ in (8.21).

The above three cases cover all possible values of \mathbf{p} and \mathbf{q} . This completes the proof. \square

8.1. Completing the proof of Theorem 2.6. The upper bound of Theorem 2.6 follows from Lemma 8.2, Lemma 8.3(b) and Proposition 7.7(a). The lower bound of Theorem 2.6 follows from Lemma 8.4 and Proposition 7.7(a). \square

9. Proofs of auxiliary lemmas. In this section we prove the lemmas in Section 7. Specifically, in Section 9.1 we prove Lemma 7.1, in Section 9.2 we prove Lemmas 7.2 and 7.3, in Section 9.3 we prove Lemma 7.4, in Section 9.4 we prove Lemma 7.5, and finally in Section 9.5 we prove Lemma 7.6.

We start with the following remark.

REMARK 9.1. Fix $0 \leq t_1 < t_2 < \infty$, $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \mathbb{R}_+^\infty$ and $(\xi, \psi) \in \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ such that $I_{t_1, t_2}(\xi, \psi) < \infty$. Fix $\varepsilon \in (0, 1)$. Then there exists $\varphi \in \mathcal{S}_{t_2}(\xi, \psi)$ such that

$$\sum_{k=0}^{\infty} \int_{[t_1, t_2] \times [0, 1]} \ell(\varphi_k(s, y)) \, ds \, dy \leq I_{t_1, t_2}(\xi, \psi) + \varepsilon.$$

Using convexity of ℓ , we can assume without loss of generality that $\varphi_k(t, y) = \rho_k(t) \times 1_{[0, r_k(\xi(t))]}(y) + 1_{[r_k(\xi(t)), 1]}(y)$ for $t \in [t_1, t_2]$, where ρ_k is some nonnegative function. From (2.15) and (2.16) we see that for a.e. $t \in [t_1, t_2]$,

$$(9.1) \quad \xi'_k(t) = -\rho_k(t)r_k(\xi(t)), \quad k \in \mathbb{N},$$

$$(9.2) \quad \psi'(t) = \sum_{k=0}^{\infty} (k-2)\rho_k(t)r_k(\xi(t)).$$

Since $(\xi, \psi) \in \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$, $\xi_0(t) = \xi_0(t_1) + \psi(t) - \psi(t_1)$ over (t_1, t_2) , namely there is no reflection over this interval. Therefore, for a.e. $t \in [t_1, t_2]$, we have $\xi'_0(t) = \psi'(t)$ and

$$(9.3) \quad -\frac{1}{2} \frac{d}{dt} r(\xi(t)) = -\frac{1}{2} \left(\psi'(t) + \sum_{k=1}^{\infty} k \xi'_k(t) \right) = \sum_{k=0}^{\infty} \rho_k(t)r_k(\xi(t)), \quad t \in [t_1, t_2].$$

Now we prove the lemmas in Section 7.

9.1. Proof of Lemma 7.1. We first prove (7.10). Since

$$\inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \leq I_{t_1, t_1+\xi}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \leq I_{t_1, t_1+\xi}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}),$$

it suffices to show

$$(9.4) \quad \inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \geq I_{t_1, t_1+\xi}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$$

when $\inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) < \infty$. Fix $\varepsilon \in (0, 1)$. There exist $t_2^\varepsilon \geq t_1$, $(\xi, \psi) \in \mathcal{J}_{t_1, t_2^\varepsilon}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ and $\varphi \in \mathcal{S}_{t_2^\varepsilon}(\xi, \psi)$ such that

$$(9.5) \quad \sum_{k=0}^{\infty} \int_{[t_1, t_2^\varepsilon] \times [0, 1]} \ell(\varphi_k(s, y)) \, ds \, dy \leq I_{t_1, t_2^\varepsilon}(\xi, \psi) + \varepsilon \leq \inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) + 2\varepsilon.$$

Recall that $t \mapsto r(\xi(t))$ is a nonincreasing function (see (9.3)). We claim that in fact we can assume without loss of generality that $t \mapsto r(\xi(t))$ is strictly decreasing for $t \in [t_1, t_2^\varepsilon]$. Indeed, if this function is not strictly decreasing, we can modify (ξ, ψ) such that for the modified trajectory strict monotonicity holds and the associated cost is not any higher. Such a modification can be constructed via a limiting argument as follows. Consider (ξ^n, ψ^n) defined recursively as: $(\xi^0, \psi^0, \varphi^0) \doteq (\xi, \psi, \varphi)$ on $[0, t_2^0]$ where $t_2^0 \doteq t_2^\varepsilon$. For $n \in \mathbb{N}_0$, having defined (ξ^n, ψ^n) and $t_2^n \leq t_2^\varepsilon$, where $(\xi^n, \psi^n) \in \mathcal{J}_{t_1, t_2^n}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$, and $\varphi^n \in \mathcal{S}_{t_2^n}(\xi^n, \psi^n)$ such that

$$(9.6) \quad \sum_{k=0}^{\infty} \int_{[t_1, t_2^n] \times [0, 1]} \ell(\varphi_k^n(s, y)) \, ds \, dy \leq \sum_{k=0}^{\infty} \int_{[t_1, t_2^\varepsilon] \times [0, 1]} \ell(\varphi_k(s, y)) \, ds \, dy,$$

we modify (ξ^n, ψ^n) , in case $r(\xi(\cdot))$ is not strictly decreasing on $[t_1, t_2^n]$ as follows. Let $[s_1^n, s_2^n] \subset [t_1, t_2^n]$ be the largest constant piece of $r(\xi^n(\cdot))$, namely $r(\xi^n(t))$ is constant on $[s_1^n, s_2^n]$ and $s_2^n - s_1^n$ is maximized among all such possible pieces. Let $t_2^{n+1} \doteq t_2^n - (s_2^n - s_1^n)$ and define (ξ^{n+1}, ψ^{n+1}) by shrinking (ξ^n, ψ^n) over $[s_1^n, s_2^n]$, namely let $(\xi^{n+1}(t), \psi^{n+1}(t)) \doteq (\xi^n(t), \psi^n(t))$ for $t \leq s_1^n$ and $(\xi^{n+1}(t), \psi^{n+1}(t)) \doteq (\xi^n(t + s_2^n - s_1^n), \psi^n(t + s_2^n - s_1^n))$ for $s_1^n < t \leq t_2^{n+1}$. Clearly $(\xi^{n+1}, \psi^{n+1}) \in \mathcal{J}_{t_1, t_2^{n+1}}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ and the associated control φ^{n+1} satisfies (9.6) with n replaced with $n+1$. If $r(\xi(t))$ only has N constant pieces over $[t_1, t_2^\varepsilon]$, then (ξ^N, ψ^N) is the desired modification of (ξ, ψ) . If $r(\xi(t))$ has countably many constant pieces over $[t_1, t_2^\varepsilon]$, then the sequence (ξ^n, ψ^n) is well defined and (9.6) holds for every n . Since the sequence t_2^n is nonincreasing, it converges to some point \bar{t}_2 . Since $I_{\bar{t}_2}$ has compact sublevel sets, this sequence (of paths over the time interval $[t_1, \bar{t}_2]$) has a limit point $(\bar{\xi}, \bar{\psi})$. It is easy to check that this limit point must belong to $\mathcal{J}_{t_1, \bar{t}_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ and $I_{t_1, \bar{t}_2}(\bar{\xi}, \bar{\psi}) \leq \liminf_{n \rightarrow \infty} I_{t_1, \bar{t}_2}(\xi^n, \psi^n) \leq I_{t_1, t_2^\varepsilon}(\xi, \psi)$. From the construction one can show that for fixed $\delta \in (0, 1)$, $\inf_{s \in [t_1, t_2^\varepsilon] - \delta} |r(\xi^n(s)) - r(\xi^n(s + \delta))|$ is nondecreasing in n and eventually positive. Therefore, $|r(\bar{\xi}(s)) - r(\bar{\xi}(s + \delta))| > 0$ for each $s \in [t_1, \bar{t}_2 - \delta]$. As δ is arbitrary, $(\bar{\xi}, \bar{\psi})$ is the desired modification of (ξ, ψ) verifying the claim.

From Remark 9.1 we can further assume without loss of generality that $\varphi_k(t, y) = \rho_k(t)1_{[0, r_k(\xi(t))]}(y) + 1_{[r_k(\xi(t)), 1]}(y)$ for $t \in [t_1, t_2^\varepsilon]$ and (9.1)–(9.3) hold for $t \in [t_1, t_2^\varepsilon]$. We now introduce a time transformation. Consider the nondecreasing function f defined as: $f(0) = 0$,

$$f'(t) \doteq \begin{cases} 1, & t \in [0, t_1), \\ -\frac{1}{2} \frac{d}{dt} r(\xi(t)) = \sum_{k=0}^{\infty} \rho_k(t) r_k(\xi(t)), & t \in [t_1, t_2^\varepsilon], \end{cases}$$

where the equality in the second line follows from (9.3). Since $r(\xi(t))$ is strictly decreasing for $t \in [t_1, t_2^\varepsilon]$, $f(t)$ must be strictly increasing for $t \in [t_1, t_2^\varepsilon]$. So $g \doteq f^{-1}$ is well defined and absolutely continuous on $[0, f(t_2^\varepsilon)]$. Note that $f(t_2^\varepsilon) = f(t_1) + \int_{t_1}^{t_2^\varepsilon} f'(t) \, dt = t_1 - \frac{1}{2}(r(\xi(t_2^\varepsilon)) - r(\xi(t_1))) = t_1 + \varsigma$, where the last equality is from (7.9). Define $(\tilde{\xi}(t), \tilde{\psi}(t)) \doteq (\xi(g(t)), \psi(g(t)))$ for $t \in [0, t_1 + \varsigma]$. Then it is easy to see that $(\tilde{\xi}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_1 + \varsigma}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$. Since $f(g(t)) = t$, $f'(g(t))g'(t) = 1$ and so

$$\frac{d}{dt} r(\tilde{\xi}(t)) = -2f'(g(t))g'(t) = -2 \quad \text{for a.e. } t \in [t_1, t_1 + \varsigma].$$

Therefore, $(\tilde{\xi}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_1 + \varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$. Define

$$\tilde{\varphi}_k(t, y) \doteq \begin{cases} \varphi_k(t, y), & t \in [0, t_1), \\ \tilde{\rho}_k(t)1_{[0, r_k(\tilde{\xi}(t))]}(y) + 1_{[r_k(\tilde{\xi}(t)), 1]}(y), & t \in [t_1, t_1 + \varsigma], \end{cases}$$

where $\tilde{\rho}_k(t) \doteq \rho_k(g(t))g'(t)$ for $t \in [t_1, t_1 + \varsigma]$. From (9.1) and (9.2), for $t \in [t_1, t_1 + \varsigma]$,

$$\tilde{\zeta}'_k(t) = \zeta'_k(g(t))g'(t) = -\rho_k(g(t))r_k(\zeta(g(t)))g'(t) = -\tilde{\rho}_k(t)r_k(\tilde{\zeta}(t)), \quad k \in \mathbb{N},$$

$$\tilde{\psi}'(t) = \psi'(g(t))g'(t) = \sum_{k=0}^{\infty} (k-2)\rho_k(g(t))r_k(\zeta(g(t)))g'(t) = \sum_{k=0}^{\infty} (k-2)\tilde{\rho}_k(t)r_k(\tilde{\zeta}(t)).$$

So $\tilde{\varphi} \in \mathcal{S}_{t_1+\varsigma}(\tilde{\zeta}, \tilde{\psi})$. We claim that

$$(9.7) \quad \sum_{k=0}^{\infty} \int_{[t_1, t_2^{\varepsilon}] \times [0, 1]} \ell(\varphi_k(t, y)) dt dy \geq \sum_{k=0}^{\infty} \int_{[t_1, t_1+\varsigma] \times [0, 1]} \ell(\tilde{\varphi}_k(t, y)) dt dy.$$

To see the claim, first note that the left-hand side of (9.7) equals

$$\sum_{k=0}^{\infty} \int_{t_1}^{t_2^{\varepsilon}} r_k(\zeta(t))\ell(\rho_k(t)) dt.$$

Since $g(f(t)) = t$, we have $g'(f(t))f'(t) = 1$ and hence the right-hand side of (9.7) is

$$\begin{aligned} \sum_{k=0}^{\infty} \int_{t_1}^{t_1+\varsigma} r_k(\tilde{\zeta}(t))\ell(\tilde{\rho}_k(t)) dt &= \sum_{k=0}^{\infty} \int_{t_1}^{t_2^{\varepsilon}} r_k(\tilde{\zeta}(f(t)))\ell(\tilde{\rho}_k(f(t)))f'(t) dt \\ &= \sum_{k=0}^{\infty} \int_{t_1}^{t_2^{\varepsilon}} r_k(\zeta(t))\ell\left(\frac{\rho_k(t)}{f'(t)}\right)f'(t) dt. \end{aligned}$$

Combining the above two facts, we have

$$\begin{aligned} &\sum_{k=0}^{\infty} \int_{[t_1, t_2^{\varepsilon}] \times [0, 1]} \ell(\varphi_k(t, y)) dt dy - \sum_{k=0}^{\infty} \int_{[t_1, t_1+\varsigma] \times [0, 1]} \ell(\tilde{\varphi}_k(t, y)) dt dy \\ &= \sum_{k=0}^{\infty} \int_{t_1}^{t_2^{\varepsilon}} r_k(\zeta(t))[\rho_k(t) \log f'(t) - f'(t) + 1] dt \\ &= \int_{t_1}^{t_2^{\varepsilon}} \left[\left(\sum_{k=0}^{\infty} r_k(\zeta(t))\rho_k(t) \right) \log f'(t) - f'(t) + 1 \right] dt, \\ &= \int_{t_1}^{t_2^{\varepsilon}} \ell(f'(t)) dt \geq 0, \end{aligned}$$

where the next to last equality uses the fact that $\sum_{k=0}^{\infty} r_k(\zeta(t)) = 1$ for all $t \in [t_1, t_2^{\varepsilon}]$ and the last equality uses the definition of $f'(t)$. This proves the claim in (9.7). Combining (9.5) and (9.7) with the fact that $\tilde{\varphi} \in \mathcal{S}_{t_1+\varsigma}(\tilde{\zeta}, \tilde{\psi})$ and $(\tilde{\zeta}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ gives

$$I_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \leq \sum_{k=0}^{\infty} \int_{[t_1, t_1+\varsigma] \times [0, 1]} \ell(\tilde{\varphi}_k(t, y)) dt dy \leq \inf_{t_2 \geq t_1} I_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) + 2\varepsilon.$$

Since $\varepsilon \in (0, 1)$ is arbitrary, (9.4) follows, which, as argued previously, gives (7.10).

Next consider (7.11) and the third statement in the lemma for fixed $(\zeta, \psi) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$. We first show that

$$(9.8) \quad I_{t_1, t_1+\varsigma}(\zeta, \psi) \geq \int_{t_1}^{t_1+\varsigma} L(\zeta(s), \zeta'(s)) ds.$$

Assume without loss of generality that $I_{t_1, t_1 + \varsigma}(\xi, \psi) < \infty$. Fix $\varepsilon \in (0, 1)$. From Remark 9.1 we can find some $\varphi \in \mathcal{S}_{t_1 + \varsigma}(\xi, \psi)$ such that

$$\sum_{k=0}^{\infty} \int_{[t_1, t_1 + \varsigma] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy \leq I_{t_1, t_1 + \varsigma}(\xi, \psi) + \varepsilon,$$

$$\varphi_k(t, y) = \rho_k(t) 1_{[0, r_k(\xi(t))]}(y) + 1_{[r_k(\xi(t)), 1]}(y), \quad t \in [t_1, t_1 + \varsigma], k \in \mathbb{N}_0,$$

for a suitable sequence of nonnegative functions ρ_k , and (9.1)–(9.3) hold for $t \in [t_1, t_1 + \varsigma]$. Using (9.1), (9.2) and the fact that $\frac{d}{dt}r(\xi(t)) = -2$ for a.e. $t \in [t_1, t_1 + \varsigma]$ we have

$$\rho_0(t)r_0(\xi(t)) = -\frac{\psi'(t) + \sum_{k=1}^{\infty} (k-2)\zeta'_k(t)}{2} = 1 + \sum_{k=1}^{\infty} \zeta'_k(t),$$

which also implies $\sum_{k=1}^{\infty} \zeta'_k(t) \geq -1$ for a.e. $t \in [t_1, t_1 + \varsigma]$, proving the third statement in the lemma. Furthermore we have

$$\begin{aligned} (9.9) \quad & \sum_{k=0}^{\infty} \int_{[t_1, t_1 + \varsigma] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy \\ &= \sum_{k=0}^{\infty} \int_{t_1}^{t_1 + \varsigma} r_k(\xi(t)) \ell(\rho_k(t)) dt \\ &= \int_{t_1}^{t_1 + \varsigma} \left[r_0(\xi(t)) \ell\left(\frac{1 + \sum_{k=1}^{\infty} \zeta'_k(t)}{r_0(\xi(t))}\right) + \sum_{k=1}^{\infty} r_k(\xi(t)) \ell\left(\frac{-\zeta'_k(t)}{r_k(\xi(t))}\right) \right] dt \\ &= \int_{t_1}^{t_1 + \varsigma} L(\xi(t), \xi'(t)) dt, \end{aligned}$$

where the last equality uses the definition of ℓ in (2.14) and L in (7.6) and we use the convention that $0\ell(x/0) = 0$ for $x \geq 0$. Therefore,

$$(9.10) \quad \int_{t_1}^{t_1 + \varsigma} L(\xi(t), \xi'(t)) dt = \sum_{k=0}^{\infty} \int_{[t_1, t_1 + \varsigma] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy \leq I_{t_1, t_1 + \varsigma}(\xi, \psi) + \varepsilon.$$

Since $\varepsilon \in (0, 1)$ is arbitrary, we have (9.8).

Next we show that

$$(9.11) \quad I_{t_1, t_1 + \varsigma}(\xi, \psi) \leq \int_{t_1}^{t_1 + \varsigma} L(\xi(s), \xi'(s)) ds.$$

Assume without loss of generality that $\int_{t_1}^{t_1 + \varsigma} L(\xi(s), \xi'(s)) ds < \infty$. Since there exists some $(\xi^*, \psi^*) \in \mathcal{J}_{0, t_1}^0(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})$ such that $I_{0, t_1}(\xi^*, \psi^*) < \infty$, we can further assume without loss of generality that $I_{0, t_1}(\xi, \psi) < \infty$. Then there exists some $\varphi^* \in \mathcal{S}_{t_1}(\xi, \psi)$. Let $\varphi(t, y) \doteq \varphi^*(t, y)$ for $t \in [0, t_1]$, and for $t \in [t_1, t_1 + \varsigma]$ define

$$\rho_k(t) \doteq -\frac{\zeta'_k(t)}{r_k(\xi(t))} \mathbf{1}_{\{r_k(\xi(t)) \neq 0\}}, \quad k \in \mathbb{N},$$

$$\rho_0(t) \doteq \frac{\sum_{k=1}^{\infty} (k-2)\rho_k(t)r_k(\xi(t)) - \psi'(t)}{2r_0(\xi(t))} \mathbf{1}_{\{r_0(\xi(t)) \neq 0\}},$$

$$\varphi_k(t, y) \doteq \rho_k(t) 1_{[0, r_k(\xi(t))]}(y) + 1_{[r_k(\xi(t)), 1]}(y), \quad y \in [0, 1], k \in \mathbb{N}_0.$$

Clearly (9.1) and (9.2) hold for $t \in [t_1, t_1 + \varsigma]$ and hence $\varphi \in \mathcal{S}_{t_1 + \varsigma}(\zeta, \psi)$. Also one can check that (9.9) still holds. Therefore,

$$I_{t_1, t_1 + \varsigma}(\zeta, \psi) \leq \sum_{k=0}^{\infty} \int_{[t_1, t_1 + \varsigma] \times [0, 1]} \ell(\varphi_k(s, y)) ds dy = \int_{t_1}^{t_1 + \varsigma} L(\zeta(t), \zeta'(t)) dt.$$

This gives (9.11) and completes the proof of (7.11).

Finally, (7.12) follows on combining (7.10), (7.3) and (7.11). This completes the proof of the lemma. \square

9.2. Proofs of Lemmas 7.2 and 7.3. PROOF OF LEMMA 7.2. Consider for $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \Xi$, the function $\alpha \mapsto B(\alpha)$ on $(0, 1)$, defined by

$$\begin{aligned} B(\alpha) \doteq B(\alpha; \mathbf{x}^{(1)}, \mathbf{x}^{(2)}) &\doteq \frac{1}{\alpha} \left((1 - \alpha^2) \sum_{k=1}^{\infty} \frac{kz_k}{1 - \alpha^k} - \sum_{k=1}^{\infty} kz_k + x_0^{(2)} - \alpha^2 x_0^{(1)} \right) \\ (9.12) \quad &= z_1 - \sum_{k=3}^{\infty} kz_k \frac{\alpha - \alpha^{k-1}}{1 - \alpha^k} + \frac{x_0^{(2)}}{\alpha} - \alpha x_0^{(1)} \\ &= z_1 - \sum_{k=3}^{\infty} kz_k B_k(\alpha) + \frac{x_0^{(2)}}{\alpha} - \alpha x_0^{(1)}, \end{aligned}$$

where $B_k(\alpha) \doteq (\alpha - \alpha^{k-1})/(1 - \alpha^k)$ for $k \geq 3$ and $z = \mathbf{x}^{(1)} - \mathbf{x}^{(2)}$ as before. For each $k \geq 3$ and $\alpha \in (0, 1)$, using the inequality of arithmetic and geometric means one can verify that

$$(9.13) \quad B'_k(\alpha) = \frac{(1 - \alpha^2)(k - 1)}{(1 - \alpha^k)^2} \left(\frac{1 + \alpha^2 + \alpha^4 + \cdots + \alpha^{2k-4}}{k - 1} - \alpha^{k-2} \right) > 0,$$

and

$$(9.14) \quad 0 = B_k(0+) \leq B_k(\alpha) \leq B_k(1-) = (k - 2)/k.$$

So $B(1-) = z_1 - \sum_{k=3}^{\infty} (k - 2)z_k + x_0^{(2)} - x_0^{(1)} = -(\sum_{k=1}^{\infty} (k - 2)z_k + z_0) < 0$ by assumption and $B(\alpha)$ is decreasing in $\alpha \in (0, 1)$. Also note that the assumption $\sum_{k=1}^{\infty} kz_k + z_0 > 2 \sum_{k=1}^{\infty} z_k$ can be rewritten as

$$\sum_{k=3}^{\infty} (k - 2)z_k + x_0^{(1)} > z_1 + x_0^{(2)},$$

which implies

$$(9.15) \quad \text{either } x_0^{(1)} > 0 \text{ or } z_k > 0 \text{ for some } k \geq 3.$$

From this and (9.13) we see that $B(\alpha)$ is actually strictly decreasing in $\alpha \in (0, 1)$. Since each $B_k(\alpha)$ is continuous on $(0, 1)$, $B(\alpha)$ is also continuous by (9.14) and the dominated convergence theorem. Finally, since $B(0+) = z_1 + \infty \cdot 1_{\{x_0^{(2)} > 0\}} > 0$ and $B(1-) < 0$, there must exist a unique $\beta \in (0, 1)$ such that $B(\beta) = 0$. This completes the proof of the lemma. \square

PROOF OF LEMMA 7.3. Suppose $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \rightarrow (\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ as $n \rightarrow \infty$, where $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}), (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \Xi$. Recall the function $B(\cdot)$ defined above (9.13) and the definition of $\beta(\cdot)$ from Section 2.2. We consider two possible cases for the values of $x_0^{(2)}$ and z_1 .

Case 1: $x_0^{(2)} = 0$ and $z_1 = 0$. In this case $\beta = \beta(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = 0$ by definition and $x_0^{(2)} \log \beta = 0$ by our convention. Since $B(0+) = z_1 + \infty \cdot 1_{\{x_0^{(2)} > 0\}} = 0$ and $B(\alpha)$ is strictly decreasing in $\alpha \in (0, 1)$, we have $B(\alpha) < 0$ for every $\alpha \in (0, 1)$. Fixing $\alpha \in (0, 1)$, from (9.14) and the dominated convergence theorem one has

$$(9.16) \quad B^{(n)}(\alpha) \doteq B(\alpha; \mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \rightarrow B(\alpha) \doteq B(\alpha; \mathbf{x}^{(1)}, \mathbf{x}^{(2)})$$

as $n \rightarrow \infty$. Therefore, $B^{(n)}(\alpha) < 0$ for sufficiently large n . Since $B^{(n)}$ is decreasing, we must have $\beta^{(n)} \doteq \beta(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) \leq \alpha$ for all such n . Since $\alpha \in (0, 1)$ is arbitrary, this implies that as $n \rightarrow \infty$, $\beta^{(n)} \rightarrow 0 = \beta$. Next note that the convergence of $x_0^{(2),n} \log \beta^{(n)} \rightarrow x_0^{(2)} \log \beta = 0$ holds trivially if $x_0^{(2),n} = 0$ for all sufficiently large n . Suppose now that $x_0^{(2),n} > 0$ for every n . Also take n to be sufficiently large, so that $x_0^{(2),n} < 1$. From (9.12) and since $kB_k^{(n)}(\alpha) \leq (k-2)$ from (9.14) [applied with $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ replaced by $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})$] we have

$$B^{(n)}((x_0^{(2),n})^2) \geq 0 - \sum_{k=3}^{\infty} (k-2)z_k^n + \frac{1}{x_0^{(2),n}} - (x_0^{(2),n})^2 x_0^{(1),n} > 0$$

for $x_0^{(2),n}$ sufficiently small. So $\beta^{(n)} \geq (x_0^{(2),n})^2$ for such n and so $x_0^{(2),n} \log \beta^{(n)} \rightarrow 0 = x_0^{(2)} \log \beta$.

Case 2: $x_0^{(2)} > 0$ or $z_1 > 0$. In this case, for n sufficiently large, we must have $x_0^{(2),n} > 0$ or $z_1^n > 0$. So $\beta^{(n)}$ satisfies $B^{(n)}(\beta^{(n)}) = 0$ for all such n . Since $\beta > 0$, and by proof of Lemma 7.2 $B(\beta) = 0$, $B(0+) > 0$ and $B(\cdot)$ is strictly decreasing, we have $B(\beta/2) > 0$. As in the proof of (9.16), we see that $B^{(n)}(\beta/2) \rightarrow B(\beta/2)$ as $n \rightarrow \infty$, and so $B^{(n)}(\beta/2) > 0$ for all sufficiently large n . Since $B^{(n)}$ is decreasing, we must have $\beta^{(n)} \geq \beta/2 > 0$ for all sufficiently large n . From this, (9.14) and the dominated convergence theorem one can show that along any convergent subsequence of $\beta^{(n)}$, $B^{(n)}(\beta^{(n)}) \rightarrow B(\lim \beta^{(n)})$. So any limit point of $\beta^{(n)}$ is a solution to $B(\alpha) = 0$ defined on $(0, 1)$. But β is the unique solution to this equation. So $\beta^{(n)} \rightarrow \beta$ and also $x_0^{(2),n} \log \beta^{(n)} \rightarrow x_0^{(2)} \log \beta$. This completes the proof of the lemma. \square

9.3. Proof of Lemma 7.4. (a) Recall the definition of ς in (7.9) and $\tilde{\varsigma}, \tilde{z}_k$ from Construction 1. From (7.13) we have

$$(9.17) \quad \tilde{\varsigma} = \frac{\varsigma}{1 - \beta^2} = \frac{x_0^{(1)} - x_0^{(2)} + \sum_{k=1}^{\infty} kz_k}{2(1 - \beta^2)} = \frac{1}{2} \left(x_0^{(1)} + \sum_{k=1}^{\infty} k\tilde{z}_k \right).$$

Since $\beta \in [0, 1)$, we have $\tilde{\varsigma} \leq \varsigma/(1 - \beta^2) = \tilde{\varsigma}$. This proves part (a).

(b) We first show that $(\tilde{\varsigma}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_1 + \varsigma}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$. For this, it suffices to check

$$(9.18) \quad \tilde{\zeta}_k(t_1 + \varsigma) = p_k^{(2)} \quad \text{for } k \in \mathbb{N}, \quad \tilde{\zeta}_0(t_1 + \varsigma) = x_0^{(2)},$$

$$(9.19) \quad \tilde{\psi}(t) - \tilde{\psi}(t_1) \geq -x_0^{(1)} \quad \text{for } t \in [t_1, t_1 + \varsigma].$$

From (7.14) we have

$$\tilde{\zeta}_k(t_1 + \varsigma) = p_k^{(1)} - \tilde{z}_k(1 - (1 - \varsigma/\tilde{\varsigma})^{k/2}) = p_k^{(1)} - \tilde{z}_k(1 - \beta^k) = p_k^{(1)} - z_k = p_k^{(2)},$$

which gives the first statement in (9.18). From this, (7.15) and (7.9) it follows that

$$\tilde{\zeta}_0(t_1 + \varsigma) = x_0^{(1)} + \sum_{k=1}^{\infty} k(p_k^{(1)} - p_k^{(2)}) - 2\varsigma = x_0^{(2)},$$

which gives the second statement in (9.18).

For (9.19), applying the change of variable $t - t_1 = \tilde{\zeta}(1 - \alpha_t^2)$, namely $\alpha_t = (1 - \frac{t-t_1}{\tilde{\zeta}})^{1/2}$ for $t \in [t_1, t_1 + \varsigma]$, we have

$$\begin{aligned} \sum_{k=1}^{\infty} k(p_k^{(1)} - \tilde{\zeta}_k(t)) - 2(t - t_1) &= \sum_{k=1}^{\infty} k\tilde{z}_k \left[1 - \left(1 - \frac{t-t_1}{\tilde{\zeta}}\right)^{k/2} \right] - 2(t - t_1) \\ &= \sum_{k=1}^{\infty} k\tilde{z}_k(1 - \alpha_t^k) - 2\tilde{\zeta}(1 - \alpha_t^2) \\ &= \sum_{k=1}^{\infty} k\tilde{z}_k(1 - \alpha_t^k) - \sum_{k=1}^{\infty} k\tilde{z}_k(1 - \alpha_t^2) - x_0^{(1)}(1 - \alpha_t^2) \\ &\doteq F(\alpha_t), \end{aligned}$$

where the third equality follows from part (a). Using this we can write, for $t \in [t_1, t_1 + \varsigma]$,

$$(9.20) \quad \tilde{\zeta}_0(t) = x_0^{(1)} + F(\alpha_t) = \tilde{\psi}(t) - \tilde{\psi}(t_1) + x_0^{(1)}.$$

Note that, for $t \in [t_1, t_1 + \varsigma]$,

$$\begin{aligned} x_0^{(1)} + F(\alpha_t) &= \alpha_t^2 x_0^{(1)} + \sum_{k=1}^{\infty} k\tilde{z}_k(\alpha_t^2 - \alpha_t^k) \\ &= \alpha_t(1 - \alpha_t) \left(\frac{\alpha_t x_0^{(1)}}{1 - \alpha_t} - \tilde{z}_1 + \sum_{k=3}^{\infty} k\tilde{z}_k \frac{\alpha_t - \alpha_t^{k-1}}{1 - \alpha_t} \right) \\ &= \alpha_t(1 - \alpha_t) \left(\frac{\alpha_t x_0^{(1)}}{1 - \alpha_t} - \tilde{z}_1 + \sum_{k=3}^{\infty} k\tilde{z}_k \tilde{B}_k(\alpha_t) \right) \\ &= \alpha_t(1 - \alpha_t) \tilde{B}(\alpha_t), \end{aligned}$$

where $\tilde{B}_k(\alpha) \doteq (\alpha - \alpha^{k-1})/(1 - \alpha)$ for $k \geq 3$ and $\tilde{B}(\alpha) \doteq \frac{\alpha x_0^{(1)}}{1 - \alpha} - \tilde{z}_1 + \sum_{k=3}^{\infty} k\tilde{z}_k \tilde{B}_k(\alpha)$. One can verify (e.g., using Young's inequality) that

$$\tilde{B}'_k(\alpha) = \frac{(k-2)\alpha^{k-1} - (k-1)\alpha^{k-2} + 1}{(\alpha-1)^2} > 0, \quad k \geq 3, \alpha \in [0, 1].$$

So $\tilde{B}(\alpha)$ is increasing. Using (7.13) one can verify that $\tilde{B}(\beta) = \frac{x_0^{(1)}}{\beta(1-\beta)} 1_{\{\beta>0\}} \geq 0$. Since for $t \in (t_1, t_1 + \varsigma]$, $\alpha_t \in [\beta, 1]$, for all such t $x_0^{(1)} + F(\alpha_t) \geq 0$. This along with (9.20) gives (9.19).

So far we have verified that $(\tilde{\xi}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_2}^1(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$. From (7.15) we also have $\frac{d}{dt} r(\tilde{\xi}(t)) = -2$ for $t \in [t_1, t_1 + \varsigma]$. Thus actually $(\tilde{\xi}, \tilde{\psi}) \in \mathcal{J}_{t_1, t_2}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$, completing the proof of (b).

(c) Since for $t \in (t_1, t_1 + \varsigma)$, $\tilde{\zeta}_0(t) = x_0^{(1)} + F(\alpha_t) = \alpha_t(1 - \alpha_t)\tilde{B}(\alpha_t)$ and $\tilde{B}(\beta) \geq 0$, it suffices to show that $\tilde{B}(\alpha)$ is strictly increasing in $\alpha \in [\beta, 1]$. But thanks to (9.15), this is immediate from the fact that $\tilde{B}'_k(\alpha) > 0$ and $\frac{\alpha x_0^{(1)}}{1-\alpha}$ is strictly increasing when $x_0^{(1)} > 0$. This gives part (c) and completes the proof of the lemma.

9.4. *Proof of Lemma 7.5.* Let $\mu_0 \doteq \frac{1}{2}(x_0^{(1)} + \sum_{k=1}^{\infty} kp_k^{(1)})$. Recall $\tilde{z}_k = \frac{z_k}{1-\beta^k}$. It then follows from (7.14), (7.15) and Lemma 7.4(a) that for $t \in [t_1, t_1 + \varsigma]$,

$$(9.21) \quad \begin{aligned} \tilde{\zeta}'_k(t) &= -\frac{k\tilde{z}_k}{2\tilde{\zeta} - 2(t - t_1)} \left(1 - \frac{t - t_1}{\tilde{\zeta}}\right)^{k/2} \\ &= -\frac{k}{2\tilde{\zeta} - 2(t - t_1)} [\tilde{\zeta}_k(t) - p_k^{(1)} + \tilde{z}_k], \end{aligned}$$

$$(9.22) \quad 1 + \sum_{k=1}^{\infty} \tilde{\zeta}'_k(t) = \frac{x_0^{(1)} - 2(t - t_1) + \sum_{k=1}^{\infty} k[p_k^{(1)} - \tilde{\zeta}_k(t)]}{2\tilde{\zeta} - 2(t - t_1)} = \frac{\tilde{\zeta}_0(t)}{2\tilde{\zeta} - 2(t - t_1)},$$

$$r(\tilde{\zeta}(t)) = 2\mu_0 - 2(t - t_1).$$

From these we have

$$(9.23) \quad \begin{aligned} &\int_{t_1}^{t_1+\varsigma} L(\tilde{\zeta}(s), \tilde{\zeta}'(s)) ds \\ &= \int_{t_1}^{t_1+\varsigma} \left[\left(1 + \sum_{k=1}^{\infty} \tilde{\zeta}'_k(t)\right) \log\left(\frac{1 + \sum_{k=1}^{\infty} \tilde{\zeta}'_k(t)}{\tilde{\zeta}_0(t)/r(\tilde{\zeta}(t))}\right) \right. \\ &\quad \left. + \sum_{k=1}^{\infty} (-\tilde{\zeta}'_k(t)) \log\left(\frac{-\tilde{\zeta}'_k(t)}{k\tilde{\zeta}_k(t)/r(\tilde{\zeta}(t))}\right) \right] dt \\ &= \int_{t_1}^{t_1+\varsigma} \left[\log(2\mu_0 - 2(t - t_1)) - \log(2\tilde{\zeta} - 2(t - t_1)) \right. \\ &\quad \left. - \sum_{k=1}^{\infty} \tilde{\zeta}'_k(t) \log\left(\frac{\tilde{\zeta}_k(t) - p_k^{(1)} + \tilde{z}_k}{\tilde{\zeta}_k(t)}\right) \right] dt. \end{aligned}$$

We claim that we can interchange the integration and summation in the last line. To see this, first note that there exists some $M \in \mathbb{N}$ such that $z_k \leq \tilde{z}_k \leq 2z_k \leq 2p_k^{(1)} < 1$ for $k \geq M$. Since $\tilde{\zeta}_k(t)$ is nonincreasing, we have

$$\begin{aligned} &\sum_{k=M}^{\infty} \int_{t_1}^{t_1+\varsigma} \left| \tilde{\zeta}'_k(t) \log\left(\frac{\tilde{\zeta}_k(t) - p_k^{(1)} + \tilde{z}_k}{\tilde{\zeta}_k(t)}\right) \right| dt \\ &= \sum_{k=M}^{\infty} \int_{t_1}^{t_1+\varsigma} \left| \log\left(\frac{\tilde{\zeta}_k(t) - p_k^{(1)} + \tilde{z}_k}{\tilde{\zeta}_k(t)}\right) \right| d(p_k^{(1)} - \tilde{\zeta}_k(t)) \\ &= \sum_{k=M}^{\infty} \int_0^{z_k} \left| \log\left(\frac{\tilde{z}_k - u}{p_k^{(1)} - u}\right) \right| du \\ &\leq \sum_{k=M}^{\infty} \int_0^{z_k} (-\log(\tilde{z}_k - u) - \log(p_k^{(1)} - u)) du \\ &\leq -2 \sum_{k=M}^{\infty} \int_0^{2p_k^{(1)}} \log u du. \end{aligned}$$

Using $\tilde{\ell}(x) \doteq x \log x - x = \int \log x \, dx$, the last expression equals

$$-2 \sum_{k=M}^{\infty} (\tilde{\ell}(2p_k^{(1)}) - \tilde{\ell}(0)) = 4 \sum_{k=M}^{\infty} p_k^{(1)} \log\left(\frac{1}{p_k^{(1)}}\right) - 4(\log 2 - 1) \sum_{k=M}^{\infty} p_k^{(1)}.$$

Here the last term is clearly finite. Letting $\tilde{M} \doteq \sum_{k=M}^{\infty} p_k^{(1)} \in (0, 1]$, we have

$$\begin{aligned} \sum_{k=M}^{\infty} p_k^{(1)} \log \left(\frac{1}{p_k^{(1)}} \right) &= \tilde{M} \sum_{k=M}^{\infty} \frac{p_k^{(1)}}{\tilde{M}} \log \left(\frac{1}{k^2 p_k^{(1)}} \right) + \sum_{k=M}^{\infty} p_k^{(1)} \log k^2 \\ &\leq \tilde{M} \log \left(\sum_{k=M}^{\infty} \frac{1}{\tilde{M} k^2} \right) + 2 \sum_{k=M}^{\infty} k p_k^{(1)} < \infty, \end{aligned}$$

where the inequality holds since $\log x$ is concave and $\log x \leq x$. Therefore,

$$\sum_{k=M}^{\infty} \int_{t_1}^{t_1+\varsigma} \left| \tilde{\xi}'_k(t) \log \left(\frac{\tilde{\xi}_k(t) - p_k^{(1)} + \tilde{z}_k}{\tilde{\xi}_k(t)} \right) \right| dt < \infty.$$

One can easily verify that, for $1 \leq k \leq M$,

$$\int_{t_1}^{t_1+\varsigma} \left| \tilde{\xi}'_k(t) \log \left(\frac{\tilde{\xi}_k(t) - p_k^{(1)} + \tilde{z}_k}{\tilde{\xi}_k(t)} \right) \right| dt = \int_0^{\tilde{z}_k} \left| \log \left(\frac{\tilde{z}_k - u}{p_k^{(1)} - u} \right) \right| du < \infty.$$

So the claim holds. Actually we have also shown that $\int_{t_1}^{t_1+\varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds < \infty$.

From (9.23) it then follows that

$$\begin{aligned} (9.24) \quad & \int_{t_1}^{t_1+\varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds \\ &= \int_{t_1}^{t_1+\varsigma} [\log(\mu_0 - (t - t_1)) - \log(\tilde{\varsigma} - (t - t_1))] dt \\ &\quad - \sum_{k=1}^{\infty} \int_{t_1}^{t_1+\varsigma} \log \left(\frac{\tilde{\xi}_k(t) - p_k^{(1)} + \tilde{z}_k}{\tilde{\xi}_k(t)} \right) d\tilde{\xi}_k(t) \\ &= \left[-\tilde{\ell}(\mu_0 - (t - t_1)) + \tilde{\ell}(\tilde{\varsigma} - (t - t_1)) \right. \\ &\quad \left. - \sum_{k=1}^{\infty} \tilde{\ell}(\tilde{\xi}_k(t) - p_k^{(1)} + \tilde{z}_k) + \sum_{k=1}^{\infty} \tilde{\ell}(\tilde{\xi}_k(t)) \right] \Big|_{t=t_1}^{t_1+\varsigma} \\ &= -(\mu_0 - \varsigma) \log(\mu_0 - \varsigma) + (\tilde{\varsigma} - \varsigma) \log(\tilde{\varsigma} - \varsigma) + \mu_0 \log \mu_0 - \tilde{\varsigma} \log \tilde{\varsigma} \\ &\quad + \sum_{k=1}^{\infty} [-(\tilde{z}_k - z_k) \log(\tilde{z}_k - z_k) + p_k^{(2)} \log p_k^{(2)} + \tilde{z}_k \log \tilde{z}_k - p_k^{(1)} \log p_k^{(1)}], \end{aligned}$$

where the last line follows from $\tilde{\xi}(t_1) = \mathbf{x}^{(1)}$ and $\tilde{\xi}(t_1 + \varsigma) = \mathbf{x}^{(2)}$. Using $\tilde{\varsigma} = \varsigma/(1 - \beta^2)$,

$$\begin{aligned} & (\tilde{\varsigma} - \varsigma) \log(\tilde{\varsigma} - \varsigma) - \tilde{\varsigma} \log \tilde{\varsigma} \\ &= \frac{\beta^2 \varsigma}{1 - \beta^2} \log \frac{\beta^2 \varsigma}{1 - \beta^2} - \frac{\varsigma}{1 - \beta^2} \log \frac{\varsigma}{1 - \beta^2} \\ &= -\varsigma \log \varsigma + \varsigma \log(1 - \beta^2) + \frac{2\beta^2 \varsigma}{1 - \beta^2} \log \beta. \end{aligned}$$

Since $\tilde{z}_k = z_k/(1 - \beta^k)$, we have

$$\begin{aligned}
& \sum_{k=1}^{\infty} [-(\tilde{z}_k - z_k) \log(\tilde{z}_k - z_k) + \tilde{z}_k \log \tilde{z}_k] \\
&= \sum_{k=1}^{\infty} \left[z_k \log z_k - z_k \log(1 - \beta^k) - \frac{k\beta^k z_k}{1 - \beta^k} \log \beta \right] \\
&= \sum_{k=1}^{\infty} [z_k \log z_k - z_k \log(1 - \beta^k)] + \sum_{k=1}^{\infty} \left(kz_k - \frac{kz_k}{1 - \beta^k} \right) \log \beta \\
&= \sum_{k=1}^{\infty} [z_k \log z_k - z_k \log(1 - \beta^k)] + \left(x_0^{(2)} - \frac{2\beta^2 \varsigma}{1 - \beta^2} \right) \log \beta,
\end{aligned}$$

where the last line is from (7.13) and (9.17). The last two displays along with (9.24) give

$$\begin{aligned}
& \int_{t_1}^{t_1 + \varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds \\
&= -(\mu_0 - \varsigma) \log(\mu_0 - \varsigma) - \varsigma \log \varsigma + \mu_0 \log \mu_0 + \varsigma \log(1 - \beta^2) \\
(9.25) \quad &+ \sum_{k=1}^{\infty} [z_k \log z_k + p_k^{(2)} \log p_k^{(2)} - p_k^{(1)} \log p_k^{(1)}] \\
&- \sum_{k=1}^{\infty} z_k \log(1 - \beta^k) + x_0^{(2)} \log \beta \\
&= \tilde{H}(\mathbf{z}) + \tilde{H}(\mathbf{x}^{(2)}) - \tilde{H}(\mathbf{x}^{(1)}) + \tilde{K}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}).
\end{aligned}$$

Finiteness of the above follows as in Remark 2.5. This gives the first statement in the lemma.

For the lower semicontinuity, first note that $-(\mu_0 - \varsigma) \log(\mu_0 - \varsigma) - \varsigma \log \varsigma + \mu_0 \log \mu_0 + \varsigma \log(1 - \beta^2) - \sum_{k=1}^{\infty} z_k \log(1 - \beta^k) + x_0^{(2)} \log \beta$ is continuous from Lemma 7.3 and Assumption 2.2. The remaining terms in (9.25) can be written as

$$\begin{aligned}
& \sum_{k=1}^{\infty} [z_k \log z_k + p_k^{(2)} \log p_k^{(2)} - p_k^{(1)} \log p_k^{(1)}] \\
&= \sum_{k=0}^{\infty} [z_k \log z_k + p_k^{(2)} \log p_k^{(2)} - p_k^{(1)} \log p_k^{(1)}] - [z_0 \log z_0 + p_0^{(2)} \log p_0^{(2)} - p_0^{(1)} \log p_0^{(1)}] \\
&= \sum_{k=0}^{\infty} \left[z_k \log \frac{z_k}{p_k^{(1)}} \right] + \sum_{k=0}^{\infty} \left[p_k^{(2)} \log \frac{p_k^{(2)}}{p_k^{(1)}} \right] - [z_0 \log z_0 + p_0^{(2)} \log p_0^{(2)} - p_0^{(1)} \log p_0^{(1)}],
\end{aligned}$$

where $z_0 \doteq 1 - \sum_{k=1}^{\infty} z_k$, and $p_0^{(i)} \doteq 1 - \sum_{k=1}^{\infty} p_k^{(i)}$ for $i = 1, 2$. The last term in the above display is clearly a lower semicontinuous function of $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \Xi$. The lemma follows.

9.5. Proof of Lemma 7.6. We begin with a lemma that gives the statement in Lemma 7.6 under a stronger assumption.

LEMMA 9.2. *Suppose the same setting as in Lemma 7.6. Suppose in addition that: (i) $x_0^{(1)}, x_0^{(2)} > 0$, and (ii) for every $k \in \mathbb{N}$, if $p_k^{(1)} > 0$ then $p_k^{(2)} > 0$. Then (7.18) is satisfied.*

PROOF. The first equality in (7.18) is proved in Lemma 7.1. For the second equality, we need to show that $(\tilde{\zeta}, \tilde{\psi})$ is the minimizer of the function

$$\tilde{G}(\zeta, \psi) \doteq \int_{t_1}^{t_1+\varsigma} L(\zeta(s), \zeta'(s)) ds, \quad (\zeta, \psi) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}).$$

We will prove this via contradiction. First note that $\mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ is a convex set. Also using the definition of L , one can verify that $\tilde{G}(\zeta, \psi)$ is a convex function in $(\zeta, \psi) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$. Now suppose there exists some $(\zeta, \psi) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ such that $\tilde{G}(\zeta, \psi) < \tilde{G}(\tilde{\zeta}, \tilde{\psi})$. From Lemma 7.5 we have $\tilde{G}(\tilde{\zeta}, \tilde{\psi}) < \infty$. For $\varepsilon \in [0, 1]$, construct the family of paths $(\zeta^\varepsilon, \psi^\varepsilon) \doteq (1 - \varepsilon)(\tilde{\zeta}, \tilde{\psi}) + \varepsilon(\zeta, \psi)$. Letting $g(\varepsilon) \doteq \tilde{G}(\zeta^\varepsilon, \psi^\varepsilon)$, we have $g(1) = \tilde{G}(\zeta, \psi) < \tilde{G}(\tilde{\zeta}, \tilde{\psi}) = g(0)$. It follows from the convexity that g is left and right differentiable wherever it is finite. We will show that $g'_+(0) = 0$, where $g'_+(\cdot)$ is the right derivative of g . The convexity of g will then give the desired contradiction.

By convexity of g , we have $g(\varepsilon) < g(0)$ for every $\varepsilon \in (0, 1]$. From Lemma 7.4(c), assumption (i) and continuity of $\tilde{\zeta}_0$ we have

$$(9.26) \quad \delta \doteq \inf_{t \in [t_1, t_1+\varsigma]} \tilde{\zeta}_0(t) > 0.$$

From (9.22) we see

$$(9.27) \quad 1 + \sum_{k=1}^{\infty} \tilde{\zeta}'_k(t) = \frac{\tilde{\zeta}_0(t)}{2\tilde{\zeta} - 2(t - t_1)} \geq \frac{\delta}{2\tilde{\zeta}} > 0, \quad t \in [t_1, t_1 + \varsigma].$$

Now fix $0 < \varepsilon < \frac{1}{4} \wedge \delta \wedge \frac{\delta}{2\tilde{\zeta}}$. Then $\tilde{\zeta}_0^\varepsilon(t) > \frac{\delta}{2}$ for all $t \in [t_1, t_1 + \varsigma]$.

We next argue that one can assume without loss of generality that

$$(9.28) \quad \zeta_k(t) = \tilde{\zeta}_k(t) \quad \text{for all } t \in [t_1, t_1 + \varsigma] \text{ and } k \geq n_0$$

for some large enough $n_0 \in \mathbb{N}$. To show this, we define (ζ^n, ψ^n) for $n \in \mathbb{N}$ as follows: For $t \in [0, t_1]$, $(\zeta^n(t), \psi^n(t)) \doteq (\zeta^\varepsilon(t), \psi^\varepsilon(t))$, and for $t \in [t_1, t_1 + \varsigma]$,

$$\begin{aligned} \zeta_k^n(t) &\doteq \tilde{\zeta}_k(t), \quad k \geq n, \\ \zeta_k^n(t) &\doteq \zeta_k^\varepsilon(t), \quad 1 \leq k < n, \\ \zeta_0^n(t) &\doteq x_0^{(1)} + \sum_{k=1}^{\infty} k(p_k^{(1)} - \zeta_k^n(t)) - 2(t - t_1), \\ \psi^n(t) &\doteq \psi^\varepsilon(t_1) + \sum_{k=1}^{\infty} k(p_k^{(1)} - \zeta_k^n(t)) - 2(t - t_1). \end{aligned}$$

From this definition we have $(\zeta_k^n)_{k \in \mathbb{N}} \rightarrow (\zeta_k^\varepsilon)_{k \in \mathbb{N}}$ in $\mathbb{C}([0, t_1 + \varsigma] : \mathbb{R}_+^\infty)$ as $n \rightarrow \infty$. So $(\zeta_0^n, \psi^n) \rightarrow (\zeta_0^\varepsilon, \psi^\varepsilon)$ in $\mathbb{C}([0, t_1 + \varsigma] : \mathbb{R}^2)$ as $n \rightarrow \infty$. From this we see $\psi^n(t) - \psi^n(t_1) + x_0^{(1)} = \zeta_0^n(t)$ is uniformly bounded away from 0 in $t \in [t_1, t_1 + \varsigma]$ for sufficiently large n . So $\zeta^n \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ for all such n . Recall L_k and L defined in (7.6). Using the definition of ζ_k^n for $1 \leq k < n$,

$$\begin{aligned} \tilde{G}(\zeta^n, \psi^n) - \tilde{G}(\zeta^\varepsilon, \psi^\varepsilon) &= \int_{t_1}^{t_1+\varsigma} [L(\zeta^n(s), (\zeta^n(s))') - L(\zeta^\varepsilon(s), (\zeta^\varepsilon(s))')] ds \\ (9.29) \quad &= \int_{t_1}^{t_1+\varsigma} [L_0(\zeta^n(s), (\zeta^n(s))') - L_0(\zeta^\varepsilon(s), (\zeta^\varepsilon(s))')] ds \\ &\quad + \int_{t_1}^{t_1+\varsigma} \sum_{k=n}^{\infty} [L_k(\zeta^n(s), (\zeta^n(s))') - L_k(\zeta^\varepsilon(s), (\zeta^\varepsilon(s))')] ds. \end{aligned}$$

We claim that both terms on the right side converge to 0 as $n \rightarrow \infty$. To see this, note that

$$\begin{aligned} L_0(\xi^n(s), (\xi^n(s))') &= \left(1 + \sum_{k=1}^{\infty} (\xi_k^n)'(s)\right) \log \left[\left(1 + \sum_{k=1}^{\infty} (\xi_k^n)'(s)\right) / \left(\frac{\xi_0^n(s)}{r(\xi^n(s))}\right) \right] \\ &\rightarrow L_0(\xi^\varepsilon(s), (\xi^\varepsilon(s))') \end{aligned}$$

as $n \rightarrow \infty$, for each $s \in [t_1, t_1 + \varsigma]$. From (9.27) and the choice of ε we have that

$$1 \geq 1 + \sum_{k=1}^{\infty} (\xi_k^n)'(s) \geq \frac{\delta}{2\tilde{\varsigma}} - \varepsilon > 0.$$

Since $\xi^\varepsilon(s)$ and $\tilde{\xi}_0(s)$ are both bounded from above and away from 0 for all $s \in [t_1, t_1 + \varsigma]$,

$$\sup_{n \in \mathbb{N}} \sup_{s \in [t_1, t_1 + \varsigma]} |L_0(\xi^n(s), (\xi^n(s))')| < \infty.$$

The first term on the right side of (9.29) then converges to 0 as $n \rightarrow \infty$ by the dominated convergence theorem. For the second term, note that

$$L_k(\xi^n(s), (\xi^n(s))') = -(\xi_k^n)'(s) \log \left[-(\xi_k^n)'(s) / \left(\frac{k\xi_k^n(s)}{r(\xi^n(s))} \right) \right] \rightarrow L_k(\xi^\varepsilon(s), (\xi^\varepsilon(s))')$$

as $n \rightarrow \infty$, for each $s \in [t_1, t_1 + \varsigma]$. Since $\xi^n \in \mathcal{J}_{t_1, t_1 + \varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$, we have $r(\xi^n(s)) = r(\tilde{\xi}(s)) = r(\xi^\varepsilon(s))$ for each $s \in [t_1, t_1 + \varsigma]$, and hence

$$|L_k(\xi^n(s), (\xi^n(s))')| \leq |L_k(\tilde{\xi}(s), \tilde{\xi}'(s))| + |L_k(\xi^\varepsilon(s), (\xi^\varepsilon(s))')|.$$

Since $\tilde{G}(\xi^\varepsilon, \psi^\varepsilon) < \infty$ and $\tilde{G}(\tilde{\xi}, \tilde{\psi}) < \infty$, we see that the last expression is summable over $k \in \mathbb{N}$ and integrable over $s \in [t_1, t_1 + \varsigma]$. Therefore, the second term in the claim converges to 0 as $n \rightarrow \infty$ by the dominated convergence theorem. From the above claim we then have that $\tilde{G}(\xi^{n_0}, \psi^{n_0}) < \tilde{G}(\tilde{\xi}, \tilde{\psi})$ for sufficiently large n_0 . We now fix such a n_0 and, abusing notation, denote $(\xi, \psi) = (\xi^{n_0}, \psi^{n_0})$ and define $(\xi^\varepsilon, \psi^\varepsilon)$ as before, by using the new definition of (ξ, ψ) , so that (9.28) holds.

Since $(\xi, \psi) \in \mathcal{J}_{t_1, t_1 + \varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$, we have $r(\xi(t)) = x_0^{(1)} + \sum_{k=1}^{\infty} kp_k^{(1)} - 2(t - t_1)$ and $\xi_0(t) = r(\xi(t)) - \sum_{k=1}^{\infty} k\xi_k(t) = x_0^{(1)} + \sum_{k=1}^{\infty} k(p_k^{(1)} - \xi_k(t)) - 2(t - t_1)$ for $t \in [t_1, t_1 + \varsigma]$. Using the definition of L , one can write

$$\begin{aligned} \tilde{G}(\xi, \psi) &= \int_{t_1}^{t_1 + \varsigma} \left\{ \left(1 + \sum_{k=1}^{\infty} \xi'_k(t)\right) \log \left[\left(1 + \sum_{k=1}^{\infty} \xi'_k(t)\right) / \left(\frac{\xi_0(t)}{r(\xi(t))}\right) \right] \right. \\ &\quad \left. - \sum_{k=1}^{\infty} \xi'_k(t) \log \left[(-\xi'_k(t)) / \left(\frac{k\xi_k(t)}{r(\xi(t))}\right) \right] \right\} dt \\ &= \int_{t_1}^{t_1 + \varsigma} \left\{ \left(1 + \sum_{k=1}^{\infty} \xi'_k(t)\right) \log \left(\frac{1 + \sum_{k=1}^{\infty} \xi'_k(t)}{x_0^{(1)} + \sum_{k=1}^{\infty} k(p_k^{(1)} - \xi_k(t)) - 2(t - t_1)} \right) \right. \\ &\quad \left. - \sum_{k=1}^{\infty} \xi'_k(t) \log \left(\frac{-\xi'_k(t)}{k\xi_k(t)} \right) \right\} dt + \int_{t_1}^{t_1 + \varsigma} \log \left(x_0^{(1)} + \sum_{k=1}^{\infty} kp_k^{(1)} - 2(t - t_1) \right) dt, \end{aligned}$$

and the analogous expression holds for $\tilde{G}(\xi^\varepsilon, \psi^\varepsilon)$. Let $\theta \doteq \xi - \tilde{\xi}$. From (9.28) we have $\theta_k = 0$ for $k > n_0$ and hence

$$\begin{aligned} g(\varepsilon) &= \int_{t_1}^{t_1+\varsigma} \left\{ \left(1 + \sum_{k=1}^{\infty} (\xi_k^\varepsilon)'(t) \right) \log \left(\frac{1 + \sum_{k=1}^{\infty} (\xi_k^\varepsilon)'(t)}{x_0^{(1)} + \sum_{k=1}^{\infty} k(p_k^{(1)} - \xi_k^\varepsilon(t)) - 2(t - t_1)} \right) \right. \\ &\quad \left. - \sum_{k=1}^{n_0} (\xi_k^\varepsilon)'(t) \log \left(\frac{-(\xi_k^\varepsilon)'(t)}{k \xi_k^\varepsilon(t)} \right) \right\} dt + C_0 \\ &= \int_{t_1}^{t_1+\varsigma} \eta(t, (\tilde{\xi}_k(t) + \varepsilon \theta_k(t), \tilde{\xi}_k'(t) + \varepsilon \theta_k'(t))_{k=1}^{n_0}) dt + C_0, \\ &\doteq \int_{t_1}^{t_1+\varsigma} \tilde{\eta}(t, \varepsilon) dt + C_0 \end{aligned}$$

for some constant C_0 , where

$$\begin{aligned} \eta(t, (u_k, v_k)_{k=1}^{n_0}) &= \left(1 + \sum_{k=1}^{n_0} v_k + \alpha_t \right) \log \left(\frac{1 + \sum_{k=1}^{n_0} v_k + \alpha_t}{x_0^{(1)} + \sum_{k=1}^{n_0} k(p_k^{(1)} - u_k) + \gamma_t - 2(t - t_1)} \right) \\ &\quad - \sum_{k=1}^{n_0} v_k \log \left(\frac{-v_k}{ku_k} \right), \end{aligned}$$

with $\alpha_t \doteq \sum_{k=n_0+1}^{\infty} \tilde{\xi}_k'(t)$ and $\gamma_t \doteq \sum_{k=n_0+1}^{\infty} k(p_k^{(1)} - \tilde{\xi}_k(t))$. We wish to show that differentiation under the integral over t with respect to ε is valid in a neighborhood of 0. For this, we now establish an integrable bound on the partial derivative of $\tilde{\eta}$ with respect to ε . To obtain such a bound, note that we only need to consider the contribution from $\varepsilon \theta_k(t)$ for $1 \leq k \leq n_0$ such that $p_k^{(2)} > 0$, since when $p_k^{(2)} = 0$, one has that $p_k^{(1)} = 0$ by assumption (ii), which implies $\theta_k(t) \equiv 0$. Therefore, assume without loss of generality that $p_k^{(2)} > 0$ for every $1 \leq k \leq n_0$. Further, note that we can assume $p_k^{(1)} > p_k^{(2)}$, since otherwise, once more, $\theta_k(t) \equiv 0$. Therefore, we assume without loss of generality that

$$(9.30) \quad p_k^{(1)} > p_k^{(2)} > 0, \quad 1 \leq k \leq n_0.$$

Denote by $\frac{\partial \eta}{\partial u_k}$ and $\frac{\partial \eta}{\partial v_k}$ the corresponding partial derivatives for the function $\eta(t, (u_k, v_k)_{k=1}^{n_0})$. Then one can verify that

$$\frac{\partial \tilde{\eta}(t, \varepsilon)}{\partial \varepsilon} = \sum_{k=1}^{n_0} \frac{\partial \eta}{\partial u_k} \Big|_{(t, (\tilde{\xi}_k^\varepsilon(t), (\tilde{\xi}_k^\varepsilon)'(t))_{k=1}^{n_0})} \theta_k(t) + \sum_{k=1}^{n_0} \frac{\partial \eta}{\partial v_k} \Big|_{(t, (\tilde{\xi}_k^\varepsilon(t), (\tilde{\xi}_k^\varepsilon)'(t))_{k=1}^{n_0})} \theta_k'(t).$$

The partial derivatives of η are

$$\begin{aligned} (9.31) \quad &\frac{\partial \eta(t, (u_k, v_k)_{k=1}^{n_0})}{\partial u_k} \\ &= \frac{k(1 + \sum_{j=1}^{n_0} v_j + \sum_{j=n_0+1}^{\infty} \tilde{\xi}_j'(t))}{x_0^{(1)} + \sum_{j=1}^{n_0} j(p_j^{(1)} - u_j) + \sum_{j=n_0+1}^{\infty} j(p_j^{(1)} - \tilde{\xi}_j(t)) - 2(t - t_1)} + \frac{v_k}{u_k}, \end{aligned}$$

$$\begin{aligned} (9.32) \quad &\frac{\partial \eta(t, (u_k, v_k)_{k=1}^{n_0})}{\partial v_k} \\ &= \log \left(\frac{1 + \sum_{j=1}^{n_0} v_j + \sum_{j=n_0+1}^{\infty} \tilde{\xi}_j'(t)}{x_0^{(1)} + \sum_{j=1}^{n_0} j(p_j^{(1)} - u_j) + \sum_{j=n_0+1}^{\infty} j(p_j^{(1)} - \tilde{\xi}_j(t)) - 2(t - t_1)} \right) \\ &\quad - \log \frac{-v_k}{ku_k}, \end{aligned}$$

for $1 \leq k \leq n_0$. For all $0 \leq \varepsilon < \frac{1}{4} \wedge \delta \wedge \frac{\delta}{2\tilde{\zeta}}$ and $t \in [t_1, t_1 + \varsigma]$, from (9.30) and (9.26),

$$0 < \frac{\delta}{2} \leq (1 - \varepsilon)\tilde{\zeta}_0(t) \leq \zeta_0^\varepsilon(t) \leq \sum_{k=1}^{\infty} kp_k < \infty,$$

$$\zeta_0^\varepsilon(t) = x_0^{(1)} + \sum_{j=1}^{n_0} j(p_j^{(1)} - \zeta_j^\varepsilon(t)) + \sum_{j=n_0+1}^{\infty} j(p_j^{(1)} - \tilde{\zeta}_j(t)) - 2(t - t_1),$$

$$0 < p_k^{(2)} \leq \zeta_k^\varepsilon(t) \leq p_k^{(1)} < \infty, \quad -1 \leq (\zeta_k^\varepsilon)'(t) \leq 0,$$

$$|\theta_k(t)| \leq p_k^{(1)}, \quad |\theta'_k(t)| \leq 2, \quad 1 \leq k \leq n_0,$$

$$0 < \frac{\delta}{4\tilde{\zeta}} \leq (1 - \varepsilon) \left(1 + \sum_{k=1}^{\infty} \tilde{\zeta}'_k(t) \right) \leq 1 + \sum_{k=1}^{\infty} (\zeta_k^\varepsilon)'(t) \leq 1,$$

where the last line uses (9.27) and Lemma 7.1. Furthermore, using (7.14) we get

$$(\zeta_k^\varepsilon)'(t) \leq (1 - \varepsilon)\tilde{\zeta}'_k(t) \leq -\frac{3k\tilde{\zeta}_k}{8\tilde{\zeta}} \left(1 - \frac{t - t_1}{\tilde{\zeta}} \right)^{k/2-1} = -\frac{3k\tilde{\zeta}_k}{8\tilde{\zeta}^{k/2}} (\tilde{\zeta} - (t - t_1))^{k/2-1}.$$

Combining these bounds we have

$$\begin{aligned} \left| \frac{\partial \eta}{\partial u_k} \right|_{(t, (\zeta_k^\varepsilon(t), (\zeta_k^\varepsilon)'(t))_{k=1}^{n_0})} &\leq \frac{k}{\delta/2} + \frac{1}{p_k^{(2)}}, \\ \left| \frac{\partial \eta}{\partial v_k} \right|_{(t, (\zeta_k^\varepsilon(t), (\zeta_k^\varepsilon)'(t))_{k=1}^{n_0})} &\leq \max \left\{ \left| \log \frac{1}{\delta/2} \right|, \left| \log \frac{\delta/4\tilde{\zeta}}{x_0^{(1)} + \sum_{j=1}^{\infty} j p_j} \right| \right\} \\ &\quad + \max \left\{ \left| \log \frac{1}{kp_k^{(2)}} \right|, \left| \log \frac{\frac{3k\tilde{\zeta}_k}{8\tilde{\zeta}^{k/2}} (\tilde{\zeta} - (t - t_1))^{k/2-1}}{kp_k^{(1)}} \right| \right\} \end{aligned}$$

for all $\varepsilon \in [0, 1/4]$, $t \in [t_1, t_1 + \varsigma]$, and $k = 1, \dots, n_0$. Therefore, one can find some $\tilde{C}_0 \in (0, \infty)$ such that

$$\left| \frac{\partial \tilde{\eta}(t, \varepsilon)}{\partial \varepsilon} \right| \leq \tilde{C}_0 + \tilde{C}_0 |\log(\tilde{\zeta} - (t - t_1))|, \quad \varepsilon \in [0, 1/4], t \in [t_1, t_1 + \varsigma].$$

Since $|\log(\tilde{\zeta} - (t - t_1))|$ is integrable in $t \in [t_1, t_1 + \varsigma]$, we have obtained an integrable bound on $|\frac{\partial \tilde{\eta}(t, \varepsilon)}{\partial \varepsilon}|$ that is uniform in $\varepsilon \in [0, 1/4]$. Thus we can differentiate under the integral sign to get

$$g'(\varepsilon) = \int_{t_1}^{t_1 + \varsigma} \frac{\partial \tilde{\eta}(t, \varepsilon)}{\partial \varepsilon} dt$$

for all $0 \leq \varepsilon < \frac{1}{4} \wedge \delta \wedge \frac{\delta}{2\tilde{\zeta}}$. Next we claim that the following Euler–Lagrange equations are satisfied.

$$(9.33) \quad \frac{\partial \eta}{\partial u_n}(t, (\tilde{\zeta}_k(t), \tilde{\zeta}'_k(t))_{k=1}^{n_0}) = \frac{d}{dt} \frac{\partial \eta}{\partial v_n}(t, (\tilde{\zeta}_k(t), \tilde{\zeta}'_k(t))_{k=1}^{n_0})$$

for $1 \leq n \leq n_0$, $t \in [t_1, t_1 + \varsigma]$.

Once this claim is verified, we have

$$\begin{aligned} g'_+(0) &= \sum_{k=1}^{n_0} \int_{t_1}^{t_1 + \varsigma} \left[\frac{\partial \eta}{\partial u_k} \Big|_{(t, (\tilde{\zeta}_k(t), \tilde{\zeta}'_k(t))_{k=1}^{n_0})} \theta_k(t) + \frac{\partial \eta}{\partial v_k} \Big|_{(t, (\tilde{\zeta}_k(t), \tilde{\zeta}'_k(t))_{k=1}^{n_0})} \theta'_k(t) \right] dt \\ &= \sum_{k=1}^{n_0} \int_{t_1}^{t_1 + \varsigma} \theta'_k(t) \left[- \int_{t_1}^t \frac{\partial \eta}{\partial u_k} \Big|_{(s, (\tilde{\zeta}_k(s), \tilde{\zeta}'_k(s))_{k=1}^{n_0})} ds + \frac{\partial \eta}{\partial v_k} \Big|_{(t, (\tilde{\zeta}_k(t), \tilde{\zeta}'_k(t))_{k=1}^{n_0})} \right] dt \end{aligned}$$

$$= \sum_{k=1}^{n_0} \int_{t_1}^{t_1+\varsigma} \tilde{c}_k \theta'_k(t) dt = \sum_{k=1}^{n_0} \tilde{c}_k (\theta_k(t_1 + \varsigma) - \theta_k(t_1)) = 0,$$

where the second equality follows from integration by parts, the third is a consequence of (9.33) with some suitable constants \tilde{c}_k and the last equality holds since $\theta_k(t_1) = 0 = \theta_k(t_1 + \varsigma)$. This gives the desired contradiction and shows that $(\tilde{\xi}, \tilde{\psi})$ is the minimizer.

Finally we prove the claim (9.33). Fix $1 \leq n \leq n_0$. Using (9.31) and (9.22) one can verify that

$$\begin{aligned} \frac{\partial \eta}{\partial u_n}(t, (\tilde{\xi}_k(t), \tilde{\xi}'_k(t))_{k=1}^{n_0}) &= \frac{n(1 + \sum_{k=1}^{\infty} \tilde{\xi}'_k(t))}{\tilde{\xi}_0(t)} + \frac{\tilde{\xi}'_n(t)}{\tilde{\xi}_n(t)} \\ &= \frac{n}{2\tilde{\xi} - 2(t - t_1)} + \frac{\tilde{\xi}'_n(t)}{\tilde{\xi}_n(t)} \\ &= \frac{d}{dt} \left(-\frac{n}{2} \log(\tilde{\xi} - (t - t_1)) + \log(\tilde{\xi}_n(t)) \right). \end{aligned}$$

Therefore, it suffices to show

$$(9.34) \quad -\frac{n}{2} \log(\tilde{\xi} - (t - t_1)) + \log(\tilde{\xi}_n(t)) = \frac{\partial \eta}{\partial v_n}(t, (\tilde{\xi}_k(t), \tilde{\xi}'_k(t))_{k=1}^{n_0}) + \bar{c}_n$$

for some constant \bar{c}_n . From (9.32) one has that

$$\begin{aligned} \frac{\partial \eta}{\partial v_n}(t, (\tilde{\xi}_k(t), \tilde{\xi}'_k(t))_{k=1}^{n_0}) &= \log(n\tilde{\xi}_n(t)) - \log(-\tilde{\xi}'_n(t)) + \log\left(\frac{1 + \sum_{k=1}^{\infty} \tilde{\xi}'_k(t)}{\tilde{\xi}_0(t)}\right) \\ &= \log(n\tilde{\xi}_n(t)) - \log(-\tilde{\xi}'_n(t)) - \log(2\tilde{\xi} - 2(t - t_1)) \end{aligned}$$

where the last line follows from (9.22). From this we have

$$\begin{aligned} -\frac{n}{2} \log(\tilde{\xi} - (t - t_1)) + \log(\tilde{\xi}_n(t)) - \frac{\partial \eta}{\partial v_n}(t, (\tilde{\xi}_k(t), \tilde{\xi}'_k(t))_{k=1}^{n_0}) \\ &= -\left(\frac{n}{2} - 1\right) \log(\tilde{\xi} - (t - t_1)) - \log\left(\frac{n}{2}\right) + \log(-\tilde{\xi}'_n(t)) \\ &= \log \tilde{z}_n - \frac{n}{2} \log \tilde{\xi}, \end{aligned}$$

where the last line follows from (9.21) and (7.14). Therefore, (9.34) holds with $\bar{c}_n = \log \tilde{z}_n - \frac{n}{2} \log \tilde{\xi}$ which proves (9.33). This completes the proof. \square

PROOF OF LEMMA 7.6. The first equality in (7.18) follows as before from Lemma 7.1. Lemma 9.2 shows that the second equality holds if the additional two assumptions in Lemma 9.2 are satisfied. Let $(\xi, \psi) \in \mathcal{J}_{t_1, t_1+\varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ be a trajectory such that $\int_{t_1}^{t_1+\varsigma} L(\xi(s), \xi'(s)) ds \leq \int_{t_1}^{t_1+\varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds$. It suffices to show

$$(9.35) \quad \int_{t_1}^{t_1+\varsigma} L(\xi(s), \xi'(s)) ds \geq \int_{t_1}^{t_1+\varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds.$$

We claim that we can assume:

- $\xi_0(t) > 0$ for all $t \in (t_1, t_1 + \varsigma)$,
- if $z_k > 0$ for some $k \in \mathbb{N}$, then $\xi_k(t) > 0$ for all $t \in (t_1, t_1 + \varsigma)$.

For this, note that $\tilde{\xi}$ satisfies these two properties. Letting $(\xi^\varepsilon, \psi^\varepsilon) \doteq \varepsilon(\xi, \psi) + (1 - \varepsilon)(\tilde{\xi}, \tilde{\psi})$ for $\varepsilon \in (0, 1)$ we have that $(\xi^\varepsilon, \psi^\varepsilon) \in \mathcal{J}_{t_1, t_1 + \varsigma}^2(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ and it satisfies the two claimed properties. Also, from the convexity of L we see that, it suffices to prove (9.35) with (ξ, ψ) replaced with $(\xi^\varepsilon, \psi^\varepsilon)$. Therefore, the claim holds.

Fix two sequences of time instants $t_1^{(n)} \doteq t_1 + \frac{1}{n}$ and $t_2^{(n)} \doteq t_1 + \varsigma - \frac{1}{n}$. Note that $t_2^{(n)} = t_1^{(n)} + \varsigma^{(n)}$ where $\varsigma^{(n)}$ is defined by (7.9) by replacing $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ with $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n}) = (\xi(t_1^{(n)}), \xi(t_2^{(n)}))$. Consider now the optimization problem in (7.3) associated with $I_{t_1^{(n)}, t_1^{(n)} + \varsigma^{(n)}}^2(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})$. Note that for this problem the two additional assumptions in Lemma 9.2 are satisfied. Furthermore, the assumption $\sum_{k=1}^{\infty} k z_k + z_0 > 2 \sum_{k=1}^{\infty} z_k$ in Lemma 7.6 also holds with z replaced by $z^{(n)} = \mathbf{x}^{(1),n} - \mathbf{x}^{(2),n}$, for sufficiently large n . Therefore, Lemma 9.2 can be applied with $(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ replaced with $(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})$. Let $(\tilde{\xi}^{(n)}, \tilde{\psi}^{(n)}) \in \mathcal{J}_{t_1^{(n)}, t_1^{(n)} + \varsigma^{(n)}}^2(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})$ be the corresponding minimizer and $\beta^{(n)} \doteq \beta(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})$. Then

$$\begin{aligned} \int_{t_1}^{t_1 + \varsigma} L(\xi(s), \xi'(s)) ds &= \lim_{n \rightarrow \infty} \int_{t_1^{(n)}}^{t_2^{(n)}} L(\xi(s), \xi'(s)) ds \\ &\geq \liminf_{n \rightarrow \infty} \int_{t_1^{(n)}}^{t_2^{(n)}} L(\tilde{\xi}^{(n)}(s), (\tilde{\xi}^{(n)})'(s)) ds \\ &= \liminf_{n \rightarrow \infty} [\tilde{H}(z^n) + \tilde{H}(\mathbf{x}^{(2),n}) - \tilde{H}(\mathbf{x}^{(1),n}) + \tilde{K}(\mathbf{x}^{(1),n}, \mathbf{x}^{(2),n})] \\ &\geq \tilde{H}(z) + \tilde{H}(\mathbf{x}^{(2)}) - \tilde{H}(\mathbf{x}^{(1)}) + \tilde{K}(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \\ &= \int_{t_1}^{t_1 + \varsigma} L(\tilde{\xi}(s), \tilde{\xi}'(s)) ds. \end{aligned}$$

Here the first inequality follows from Lemma 9.2 and the last three lines use Lemma 7.5. \square

10. Proof of LLN. In this section we give the proofs of Theorem 2.20 and Proposition 2.21.

PROOF OF THEOREM 2.20. (1) Assume without loss of generality that $T \geq 1$. Since $f_1(t) \leq 1$, we see from Assumption 2.18 that $r(\xi(\cdot))$ with r from (2.7) and ψ are well defined. Let $\varphi_k(s, y) = 1$ for all $k \in \mathbb{N}_0$ and $(s, y) \in [0, T] \times [0, 1]$. It suffices to show $\varphi \in \mathcal{S}_T(\xi, \psi)$ and $(\xi, \psi) \in \mathcal{C}_T$. Since $f_1(t) = F_1^{-1}(t)1_{[0,1]}(t)$, we have $\tau_\xi = 1$, where τ_ξ was defined in (2.17). Since $F_1(f_1(t)) = t$ for $t \in [0, 1]$,

$$f_1'(t) = -\frac{1}{\sum_{k=1}^{\infty} kp_k(f_1(t))^{k-1}} \quad \text{for } 0 < t < \tau_\xi \text{ and } f_1'(t) = 0 \text{ for } \tau_\xi < t < T.$$

Using this it follows that for $k \in \mathbb{N}$,

$$\zeta'_k(t) = -\frac{k \zeta_k(t)}{\sum_{j=1}^{\infty} j \zeta_j(t)} = -r_k(\xi(t)) \quad \text{for } 0 < t < \tau_\xi \text{ and } \zeta'_k(t) = 0 \text{ for } \tau_\xi < t < T.$$

From this we see that (2.16) holds and we can write

$$\psi(t) = \sum_{k=0}^{\infty} (k-2) \int_0^t r_k(\xi(s)) ds.$$

This gives (2.15) and verifies that $\varphi \in \mathcal{S}_T(\xi, \psi)$.

Next we argue that $(\xi, \psi) \in \mathcal{C}_T$. From Assumption 2.18, for $t < \tau_\xi$, as $K \rightarrow \infty$,

$$\sum_{k=K}^{\infty} |k-2| |\xi'_k(t)| \leq \sum_{k=K}^{\infty} kr_k(\xi(t)) \leq \frac{\sum_{k=K}^{\infty} k^2 p_k}{r(\xi(t))} \rightarrow 0.$$

In particular, ψ is absolutely continuous and thus property (a) of \mathcal{C}_T holds. Also, for $t < \tau_\xi$,

$$\psi'(t) = \sum_{k=1}^{\infty} (k-2)r_k(\xi(t)) = \frac{\sum_{k=1}^{\infty} k(k-2)p_k(f_1(t))^k}{r(\xi(t))} \leq \frac{f_1(t) \sum_{k=1}^{\infty} k(k-2)p_k}{r(\xi(t))} \leq 0.$$

Therefore, $\Gamma(\psi)(t) = 0 = \xi_0(t)$ for $t < \tau_\xi$. For $\tau_\xi \leq t \leq T$, clearly $\Gamma(\psi)(t) = 0 = \xi_0(t)$. So we have checked property (b) of \mathcal{C}_T . Property (c) of \mathcal{C}_T follows from the definition of ξ_k , $k \in \mathbb{N}$. Therefore, $(\xi, \psi) \in \mathcal{C}_T$ and part (1) follows.

(2) The fact that when $p_1 > 0$ there is a unique $\rho \in (0, 1)$ such that $G_1(\rho) = \rho$ is proved in [36]. Since $f_\rho(t) \leq 1$, we see from Assumption 2.18 that $r(\xi(\cdot))$ and ψ are well defined. Let $\varphi_k(s, y) = 1$ for all $k \in \mathbb{N}_0$ and $(s, y) \in [0, T] \times [0, 1]$. It suffices to show $\varphi \in \mathcal{S}_T(\xi, \psi)$ and $(\xi, \psi) \in \mathcal{C}_T$. First consider times $t < \tau$. Using the definitions of r , G_1 and τ , for $t < \tau$

$$\begin{aligned} r(\xi(t)) &= \mu - 2t - \mu\sqrt{1 - 2t/\mu}G_1(\sqrt{1 - 2t/\mu}) + \sum_{k=1}^{\infty} kp_k(1 - 2t/\mu)^{k/2} \\ &= \mu - 2t > \mu\rho^2 \geq 0. \end{aligned}$$

From this one can verify that for $t < \tau$,

$$\xi'_k(t) = -\frac{k\xi_k(t)}{\mu - 2t} = -r_k(\xi(t)).$$

Using this we see that (2.16) holds for $t < \tau$ and hence as before (2.15) holds as well. To show that $(\xi, \psi) \in \mathcal{C}_t$ for $t < \tau$, it suffices to show that $\psi(t)$ is absolutely continuous and $\xi_0(t) = \psi(t)$ for $t \in [0, \tau]$. Note that for $t < \tau$, $\sum_{k=1}^{\infty} |k-2|r_k(\xi(t))| \leq \frac{\sum_{k=1}^{\infty} k^2 p_k}{\mu - 2t}$. So from Assumption 2.18, ψ is absolutely continuous over $[0, \tau]$. Also, one can verify that for $t < \tau$,

$$\xi'_0(t) = \frac{d}{dt}r(\xi(t)) - \sum_{k=1}^{\infty} k\xi'_k(t) = -2 + \sum_{k=1}^{\infty} kr_k(\xi(t)) = \psi'(t).$$

So $\xi_0(t) = \psi(t)$ for $t < \tau$. Thus we have that $\varphi \in \mathcal{S}_t(\xi, \psi)$ and $(\xi, \psi) \in \mathcal{C}_t$ for each $t < \tau$.

We now consider $t \in [\tau, \tau_\xi]$. Since $\rho \in [0, 1)$ and $G_1(\rho) = \rho$, we have

$$\begin{aligned} 0 &= \frac{\mu(G_1(\rho) - \rho)}{\rho - 1} = \frac{1}{\rho - 1} \sum_{k=1}^{\infty} kp_k(\rho^{k-1} - \rho) = -p_1 + \rho \sum_{k=3}^{\infty} kp_k \frac{\rho^{k-2} - 1}{\rho - 1} \\ &= -p_1 + \rho \sum_{k=3}^{\infty} kp_k(\rho^{k-3} + \rho^{k-4} + \dots + 1) \\ &\geq -p_1 + \rho \sum_{k=3}^{\infty} kp_k(k-2)\rho^{k-3} \geq \sum_{k=1}^{\infty} k(k-2)p_k\rho^{k-1} \end{aligned}$$

and therefore $0 \geq \sum_{k=1}^{\infty} k(k-2)p_k\rho^k = \sum_{k=1}^{\infty} k(k-2)\xi_k(\tau)$. Namely, the assumption in part (1) is satisfied with ρ replaced by $\xi(\tau)$. Thus the proof for the case $t \in [\tau, \tau_\xi]$ is very similar to that in part (1), with $f_1(t)$ replaced by $f_\rho(t - \tau)$ and p_k replaced with $\xi_k(\tau)$, and we would like to omit the detail. This completes the proof of (2). \square

PROOF OF PROPOSITION 2.21. Suppose for $i = 1, 2$, $(\xi^{(i)}, \psi^{(i)})$ are two pairs such that $I_T(\xi^{(i)}, \psi^{(i)}) = 0$. By the definition of $I_T(\cdot)$, $(\xi^{(i)}, \psi^{(i)}) \in \mathcal{C}_T$. From Remark 6.3 we see that there exists some $\varphi^{(i)} \in \mathcal{S}_T(\xi^{(i)}, \psi^{(i)})$ whose cost equals $I_T(\xi^{(i)}, \psi^{(i)})$, namely

$$\sum_{k=0}^{\infty} \int_{[0, T] \times [0, 1]} \ell(\varphi_k^{(i)}(s, y)) \, ds \, dy = I_T(\xi^{(i)}, \psi^{(i)}) = 0.$$

Since $\ell(x) = 0$ if and only if $x = 1$, we must have $\varphi_k^{(i)}(s, y) = 1$ for a.e. $(s, y) \in [0, T] \times [0, 1]$ and $k \in \mathbb{N}_0$. Using such $\varphi^{(i)}$ with (2.15) and (2.16), we see that

$$(10.1) \quad \xi_k^{(i)}(t) = p_k - \int_0^t r_k(\xi^{(i)}(s)) \, ds, \quad k \in \mathbb{N},$$

$$(10.2) \quad \psi^{(i)}(t) = \sum_{k=0}^{\infty} (k-2) \int_0^t r_k(\xi^{(i)}(s)) \, ds.$$

Since $\xi_0^{(i)} = \Gamma(\psi^{(i)})$, for a.e. t , $(\xi_0^{(i)})'(t) \geq (\psi^{(i)})'(t) = \sum_{k=0}^{\infty} (k-2) r_k(\xi^{(i)}(t))$, and by (2.7)

$$\begin{aligned} \frac{d}{dt} r(\xi^{(i)}(t)) &= (\xi_0^{(i)})'(t) + \sum_{k=1}^{\infty} k (\xi_k^{(i)})'(t) \geq \sum_{k=0}^{\infty} (k-2) r_k(\xi^{(i)}(t)) - \sum_{k=1}^{\infty} k r_k(\xi^{(i)}(t)) \\ &= -2 \cdot 1_{\{r(\xi^{(i)}(t)) > 0\}} \geq -2. \end{aligned}$$

Consider the strictly increasing function $g^{(i)}(t)$ defined by

$$(10.3) \quad g^{(i)}(0) = 0, \quad (g^{(i)})'(t) = r(\xi^{(i)}(g^{(i)}(t))) 1_{\{g^{(i)}(t) < \tau_{\xi^{(i)}}\}} + 1_{\{g^{(i)}(t) \geq \tau_{\xi^{(i)}}\}},$$

where $\tau_{\xi^{(i)}}$ is as in (2.17). Since $\frac{d}{dt} r(\xi^{(i)}(t)) \in [-2, 0]$ and $0 \leq r(\xi^{(i)}(\cdot)) \leq r(\xi^{(i)}(0)) = \sum_{k=1}^{\infty} k p_k < \infty$, we see that $r(\xi^{(i)}(\cdot))$ is bounded and Lipschitz. Also $r(\xi^{(i)}(t)) > 0$ for $t < \tau_{\xi^{(i)}}$. So we have existence and uniqueness of the strictly increasing function $g^{(i)}(t)$ before it reaches $\tau_{\xi^{(i)}}$. The existence, uniqueness and monotonicity of $g^{(i)}(t)$ after $\tau_{\xi^{(i)}}$ is straightforward.

Define $(\tilde{\xi}^{(i)}(t), \tilde{\psi}^{(i)}(t)) \doteq (\xi^{(i)}(g^{(i)}(t)), \psi^{(i)}(g^{(i)}(t)))$. From (10.1) and (10.2) it follows that

$$\begin{aligned} \tilde{\xi}_k^{(i)}(t) &= p_k - \int_0^t k \tilde{\xi}_k^{(i)}(s) \, ds, \quad k \in \mathbb{N}, \\ \tilde{\psi}^{(i)}(t) &= \sum_{k=1}^{\infty} (k-2) \int_0^t k \tilde{\xi}_k^{(i)}(s) \, ds - 2 \int_0^t \tilde{\xi}_0^{(i)}(s) \, ds \\ &= \sum_{k=1}^{\infty} (k-2) \int_0^t k \tilde{\xi}_k^{(i)}(s) \, ds - 2 \int_0^t \Gamma(\tilde{\psi}^{(i)})(s) \, ds. \end{aligned}$$

Clearly $\tilde{\xi}_k^{(1)} = \tilde{\xi}_k^{(2)}$ for each $k \in \mathbb{N}$. Also, since Γ is Lipschitz on path space, Gronwall's inequality implies $\tilde{\psi}^{(1)} = \tilde{\psi}^{(2)}$, and hence $\tilde{\xi}_0^{(1)} = \tilde{\xi}_0^{(2)}$. Noting that (10.3) can be written as

$$g^{(i)}(0) = 0, \quad (g^{(i)})'(t) = r(\tilde{\xi}^{(i)}(t)) 1_{\{r(\tilde{\xi}^{(i)}(t)) > 0\}} + 1_{\{r(\tilde{\xi}^{(i)}(t)) = 0\}},$$

we have $g^{(1)} = g^{(2)}$. Since $g^{(i)}$ is strictly increasing, its inverse function is well defined and we must have that $(\xi^{(1)}, \psi^{(1)}) = (\xi^{(2)}, \psi^{(2)})$. This completes the proof. \square

Acknowledgments. We would like to thank the two referees for a careful review of our work and for the many helpful suggestions.

Funding. The research of SB was supported in part by the NSF (DMS-1606839, DMS-1613072 and DMS-2113662) and the Army Research Office (W911NF-17-1-0010). The research of AB was supported in part by the NSF (DMS-1305120, DMS-1814894, DMS-1853968). The research of PD was supported in part by the NSF (DMS-1904992) and DARPA (W911NF-15-2-0122). The research of RW was supported in part by DARPA (W911NF-15-2-0122).

REFERENCES

- [1] ALDOUS, D. (1997). Brownian excursions, critical random graphs and the multiplicative coalescent. *Ann. Probab.* **25** 812–854. MR1434128 <https://doi.org/10.1214/aop/1024404421>
- [2] AUGERI, F. (2020). Nonlinear large deviation bounds with applications to Wigner matrices and sparse Erdős-Rényi graphs. *Ann. Probab.* **48** 2404–2448. MR4152647 <https://doi.org/10.1214/20-AOP1427>
- [3] BENDER, E. A. and CANFIELD, E. R. (1978). The asymptotic number of labeled graphs with given degree sequences. *J. Combin. Theory Ser. A* **24** 296–307. MR0505796 [https://doi.org/10.1016/0097-3165\(78\)90059-6](https://doi.org/10.1016/0097-3165(78)90059-6)
- [4] BHATTACHARYA, B. B., GANGULY, S., LUBETZKY, E. and ZHAO, Y. (2017). Upper tails and independence polynomials in random graphs. *Adv. Math.* **319** 313–347. MR3695877 <https://doi.org/10.1016/j.aim.2017.08.003>
- [5] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley Series in Probability and Statistics: Probability and Statistics. Wiley, New York. MR1700749 <https://doi.org/10.1002/9780470316962>
- [6] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316. MR0595929 [https://doi.org/10.1016/S0195-6698\(80\)80030-8](https://doi.org/10.1016/S0195-6698(80)80030-8)
- [7] BORDENAVE, C. and CAPUTO, P. (2015). Large deviations of empirical neighborhood distribution in sparse random graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626 <https://doi.org/10.1016/j.aim.2008.07.008>
- [8] BORGES, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626 <https://doi.org/10.1016/j.aim.2008.07.008>
- [9] BORGES, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382 <https://doi.org/10.4007/annals.2012.176.1.2>
- [10] BOUÉ, M. and DUPUIS, P. (1998). A variational representation for certain functionals of Brownian motion. *Ann. Probab.* **26** 1641–1659. MR1675051 <https://doi.org/10.1214/aop/1022855876>
- [11] BUDHIRAJA, A., CHEN, J. and DUPUIS, P. (2013). Large deviations for stochastic partial differential equations driven by a Poisson random measure. *Stochastic Process. Appl.* **123** 523–560. MR3003362 <https://doi.org/10.1016/j.spa.2012.09.010>
- [12] BUDHIRAJA, A. and DUPUIS, P. (2000). A variational representation for positive functionals of infinite dimensional Brownian motion. *Probab. Math. Statist.* **20** 39–61. MR1785237
- [13] BUDHIRAJA, A. and DUPUIS, P. (2019). *Analysis and Approximation of Rare Events: Representations and Weak Convergence Methods. Probability Theory and Stochastic Modelling* **94**. Springer, New York. MR3967100 <https://doi.org/10.1007/978-1-4939-9579-0>
- [14] BUDHIRAJA, A., DUPUIS, P. and GANGULY, A. (2016). Moderate deviation principles for stochastic differential equations with jumps. *Ann. Probab.* **44** 1723–1775. MR3502593 <https://doi.org/10.1214/15-AOP1007>
- [15] BUDHIRAJA, A., DUPUIS, P. and MAROULAS, V. (2011). Variational representations for continuous time processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 725–747. MR2841073 <https://doi.org/10.1214/10-AIHP382>
- [16] BUDHIRAJA, A. and WU, R. (2017). Moderate deviation principles for weakly interacting particle systems. *Probab. Theory Related Fields* **168** 721–771. MR3663630 <https://doi.org/10.1007/s00440-016-0723-3>
- [17] CHATTERJEE, S. (2016). An introduction to large deviations for random graphs. *Bull. Amer. Math. Soc. (N.S.)* **53** 617–642. MR3544262 <https://doi.org/10.1090/bull/1539>
- [18] CHATTERJEE, S. (2017). *Large Deviations for Random Graphs. Lecture Notes in Math.* **2197**. Springer, Cham. MR3700183 <https://doi.org/10.1007/978-3-319-65816-2>
- [19] CHATTERJEE, S. and DEMBO, A. (2016). Nonlinear large deviations. *Adv. Math.* **299** 396–450. MR3519474 <https://doi.org/10.1016/j.aim.2016.05.017>

- [20] CHATTERJEE, S. and VARADHAN, S. R. S. (2011). The large deviation principle for the Erdős-Rényi random graph. *European J. Combin.* **32** 1000–1017. MR2825532 <https://doi.org/10.1016/j.ejc.2011.03.014>
- [21] CHOI, J. and SETHURAMAN, S. (2013). Large deviations for the degree structure in preferential attachment schemes. *Ann. Appl. Probab.* **23** 722–763. MR3059274 <https://doi.org/10.1214/12-AAP854>
- [22] COOK, N. and DEMBO, A. (2020). Large deviations of subgraph counts for sparse Erdős-Rényi graphs. *Adv. Math.* **373** 107289, 53. MR4130460 <https://doi.org/10.1016/j.aim.2020.107289>
- [23] COOK, N. A., DEMBO, A. and PHAM, H. T. (2021). Regularity method and large deviation principles for the Erdős-Rényi hypergraph. Preprint. Available at [arXiv:2102.09100](https://arxiv.org/abs/2102.09100).
- [24] DHARA, S. and SEN, S. (2019). Large deviation for uniform graphs with given degrees. Preprint. Available at [arXiv:1904.07666](https://arxiv.org/abs/1904.07666).
- [25] DUPUIS, P. and ELLIS, R. S. (1997). *A Weak Convergence Approach to the Theory of Large Deviations. Wiley Series in Probability and Statistics: Probability and Statistics*. Wiley, New York. MR1431744 <https://doi.org/10.1002/9781118165904>
- [26] DUPUIS, P., NUZMAN, C. and WHITING, P. (2004). Large deviation asymptotics for occupancy problems. *Ann. Probab.* **32** 2765–2818. MR2078557 <https://doi.org/10.1214/009117904000000135>
- [27] FORTUNATO, S. (2010). Community detection in graphs. *Phys. Rep.* **486** 75–174. MR2580414 <https://doi.org/10.1016/j.physrep.2009.11.002>
- [28] HAREL, M., MOUSSET, F. and SAMOTIJ, W. (2019). Upper tails via high moments and entropic stability. Preprint. Available at [arXiv:1904.08212](https://arxiv.org/abs/1904.08212).
- [29] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes. North-Holland Mathematical Library* **24**. North-Holland, Amsterdam–New York. MR0637061
- [30] JANSON, S. (2009). The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** 205–225. MR2497380 <https://doi.org/10.1017/S0963548308009644>
- [31] JANSON, S. and LUCZAK, M. J. (2009). A new approach to the giant component problem. *Random Structures Algorithms* **34** 197–216. MR2490288 <https://doi.org/10.1002/rsa.20231>
- [32] KARATZAS, I. and SHREVE, S. E. (1991). *Brownian Motion and Stochastic Calculus*, 2nd ed. *Graduate Texts in Mathematics* **113**. Springer, New York. MR1121940 <https://doi.org/10.1007/978-1-4612-0949-2>
- [33] KURTZ, T. G. (1981). *Approximation of Population Processes. CBMS-NSF Regional Conference Series in Applied Mathematics* **36**. SIAM, Philadelphia, PA. MR0610982
- [34] LOVÁSZ, L. (2012). *Large Networks and Graph Limits. American Mathematical Society Colloquium Publications* **60**. Amer. Math. Soc., Providence, RI. MR3012035 <https://doi.org/10.1090/coll/060>
- [35] LUBETZKY, E. and ZHAO, Y. (2017). On the variational problem for upper tails in sparse random graphs. *Random Structures Algorithms* **50** 420–436. MR3632418 <https://doi.org/10.1002/rsa.20658>
- [36] MOLLOY, M. and REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures Algorithms* **6** 161–180.
- [37] MOLLOY, M. and REED, B. (1998). The size of the giant component of a random graph with a given degree sequence. *Combin. Probab. Comput.* **7** 295–305. MR1664335 <https://doi.org/10.1017/S0963548398003526>
- [38] NEWMAN, M. E., STROGATZ, S. H. and WATTS, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64** 026118.
- [39] NEWMAN, M. E. J. (2002). Spread of epidemic disease on networks. *Phys. Rev. E* (3) **66** 016128, 11. MR1919737 <https://doi.org/10.1103/PhysRevE.66.016128>
- [40] NEWMAN, M. E. (2006). Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **103** 8577–8582.
- [41] O’CONNELL, N. (1998). Some large deviation results for sparse random graphs. *Probab. Theory Related Fields* **110** 277–285. MR1616567 <https://doi.org/10.1007/s004400050149>
- [42] PUHALSKII, A. A. (2005). Stochastic processes in random graphs. *Ann. Probab.* **33** 337–412. MR2118868 <https://doi.org/10.1214/009117904000000784>
- [43] PUHALSKII, A. A. (2013). On the number of isolated vertices in a growing random graph. *Rocky Mountain J. Math.* **43** 1941–1989. MR3178451 <https://doi.org/10.1216/RMJ-2013-43-6-1941>
- [44] VAN DER HOFSTAD, R. (2016). *Random Graphs and Complex Networks* **1**. Cambridge Univ. Press, Cambridge.