Markov Processes Relat. Fields **28**, 365–398 (2022)

Markov Processes and Related Fields

On the Metastability of a Loss Network with Diminishing Rates

Anatolii A. Puhalskii
Institute for Information Transmission Problems (IITP). 19 B. Karetny per., 127051 Moscow, Russia. E-mail: puhalski@iitp.ru

Received January 10, 2022

Abstract. A trajectorial large deviation principle is established in a mean field thermodynamic limit for a multiclass loss network with diminishing rates, which may have several stable equilibria. The large deviation limit is identified as a solution to a maxingale problem with a Markov property. The invariant measure of the network process obeys a large deviation principle as well. The network is metastable in that it spends exponentially long periods of time in the neighbourhoods of stable equilibria. A specific case of a two–class network with two stable equilibria and one unstable equilibrium is examined.

KEYWORDS: large deviations; stochastic networks; metastability; invariant measures

AMS SUBJECT CLASSIFICATION: 60F10, 60F17

1. Introduction

The following model of a cellular network was studied in Antunes et al. [2]. There are n nodes of capacity C each. Customers of K classes arrive at the nodes according to Poisson processes of respective rates α_k , $1 \leq k \leq K$. On arrival at a node, a class k customer occupies A_k units of the node's capacity, being rejected and removed from the network if the required capacity is not available. On acceptance, the customer stays at the node for an exponentially distributed length of time with mean $1/\gamma_k$ and then moves to another node, a destination being chosen uniformly at random. As on arrival, rejection occurs within the network when less than A_k units of unused capacity are available at a destination node. A class k customer may also leave the network after an exponentially distributed length of time of mean $1/\delta_k$. The arrival processes, sojourn times at the nodes, sojourn times in the network and routing decisions are independent.

Antunes et al. [2] obtained a law of large numbers for the process of the proportions of nodes with a given population, as the number of nodes goes to infinity. They analysed stability properties of the limit dynamical system and showed that it may have several stable equilibria. In Tibi [12], it was argued that the network process would spend exponentially long periods of time in the neighbourhoods of stable equilibria, in analogy with the developments in Freidlin and Wentzell [6], implying that the network is metastable. Unfortunately, the analysis in Tibi [12] is not complete.

As observed in Tibi [12], this model stands out because mean field behaviour arises in the limit only, the interactions within the network being local. Usually, when mean–field models are considered, the mean-field interaction is built in the hypotheses. In a similar vein, in the available literature multistability of a dynamical system, resulting in metastability, is assumed extraneously, for the most part, whereas in this model it is an intrinsic feature, too.

To elucidate the contribution of this paper, one needs to put things in a precise setting and review the results in Antunes et al. [2] in more detail. The state of node i at time t is described by the vector $X_i^{(n)}(t) = (X_{i,1}^{(n)}(t), \ldots, X_{i,K}^{(n)}(t))$, whose k-th entry records the number of class k customers at the node. The process $X_i^{(n)} = (X_i^{(n)}(t)\,, t \geq 0)$ takes values in the set

$$\Theta = \{\theta = (\theta_1, \ldots, \theta_K) \in \mathbb{Z}_+^K : \sum_{k=1}^{K} \theta_k A_k \leq C\}\,.$$

It is assumed that $|\Theta| \geq 2$, $|\Theta|$ denoting the cardinality of Θ. Let $Y_\theta^{(n)}(t)$ represent the proportion of nodes with θ as the population vector, i.e.,

$$Y_\theta^{(n)}(t) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{X_i^{(n)}(t)=\theta\}}$$

and let $Y^{(n)}(t) = (Y_\theta^{(n)}(t)\,, \theta \in \Theta)$, where $\mathbf{1}_\Xi$ denotes the indicator of event Ξ. As

$$\sum_{\theta\in\Theta} Y_\theta^{(n)}(t) = 1\,,$$

the process $Y^{(n)} = (Y^{(n)}(t)\,, t \geq 0)$ is a Markov process with values in the discrete simplex

$$\mathbb{S}_{|\Theta|}^{(n)} = \{y = (y_\theta,\, \theta \in \Theta) : \sum_{\theta\in\Theta} y_\theta = 1\,, y_\theta \geq 0\,, n y_\theta \text{ is an integer}\}\,.$$

Let $\mathbb{S}_{|\Theta|} = \{y = (y_\theta\,, \theta \in \Theta) : \sum_{\theta\in\Theta} y_\theta = 1\,, y_\theta \geq 0\}$. It follows from the results in Antunes et al. [2] that if the sequence $Y^{(n)}(0)$ converges in probability to $\hat{y} \in \mathbb{S}_{|\Theta|}$, as $n \to \infty$, then the sequence $Y^{(n)}$ converges in probability uniformly over compact intervals to the solution $\mathbf{y} = (\mathbf{y}(t)\,, t \geq 0)$ of the initial value problem

$$\dot{\mathbf{y}}_\theta(t) = V_\theta(\mathbf{y}(t)) \tag{1.1}$$

and $\mathbf{y}(0) = \hat{y}$, where $\mathbf{y}(t) = (\mathbf{y}_\theta(t)\,, \theta \in \Theta) \in \mathbb{S}_{|\Theta|}$ and, for $y = (y_\theta\,, \theta \in \Theta) \in \mathbb{S}_{|\Theta|}$,

$$V_\theta(y) = \sum_{k=1}^{K} \bigl((\alpha_k + \sum_{\theta' \in \Theta} \theta'_k \gamma_k y_{\theta'}) y_{\theta - e_k} + (\delta_k + \gamma_k)(\theta_k + 1) y_{\theta + e_k} - (\alpha_k + (\delta_k + \gamma_k)\theta_k + \sum_{\theta' \in \Theta} \theta'_k \gamma_k y_{\theta'}) y_\theta \bigr)\,,$$

with an overdot denoting a time derivative, e_k denoting the kth vector of the canonical basis of $\mathbb{R}^K$ and with the convention that $y_{\theta \pm e_k} = 0$ if $\theta \pm e_k \notin \Theta$.

Both $Y^{(n)}(t)$ and $\mathbf{y}(t)$ are probability distributions on Θ. The equilibrium points of (1.1) are given by an Erlang formula for the stationary distribution of an $M/M/C/C$ queue. More specifically, for $\rho = (\rho_1, \ldots, \rho_K) \in \mathbb{R}_+^K$, let a probability distribution $\nu(\rho) = (\nu_\theta(\rho)\,, \theta \in \Theta)$ on Θ be defined as

$$\nu_\theta(\rho) = \frac{1}{Z(\rho)} \prod_{k=1}^{K} \frac{\rho_k^{\theta_k}}{\theta_k!}\,, \tag{1.2}$$

$Z(\rho)$ being a normalising constant. If, for $k = 1, 2, \ldots, K$,

$$\rho_k = \frac{\alpha_k + \gamma_k \sum_{\theta \in \Theta} \theta_k \nu_\theta(\rho)}{\gamma_k + \delta_k}\,, \tag{1.3}$$

then $y = \nu(\rho)$ is an equilibrium of (1.1). Every equilibrium is of this form. The existence of solutions to (1.2) and (1.3) is proved via an application of Brouwer's fixed point theorem. On the other hand, uniqueness of an equilibrium for (1.1) might not hold and in Antunes et al. [2] an example of a network with no less than two stable equilibria is provided, so, metastability is likely to occur.

An essential stepping stone toward proving metastability is to derive a trajectorial large deviation principle (LDP) for the sequence of $Y^{(n)}$ as random elements of the associated Skorohod space. General results on large deviations of Markov processes in Freidlin and Wentzell [6] and in Wentzell [13] fall short. A major sticking point is what is known as the phenomenon of "diminishing rates", see Shwartz and Weiss [11]: near the boundary of the state space the normalised transition rates get vanishingly small, e.g., the transitions $y \to y + (f_{\theta + e_{k-1}} - f_\theta)/n$, which correspond to departures of class k customers from nodes with population vector θ, occur at the rate $n y_\theta \theta_k \delta_k$, which, when divided by n, tends to 0 as $y_\theta \to 0$, where f_θ denotes the θ-th vector of the canonical basis of $\mathbb{R}^\Theta$. The line of attack in this paper is to prove $\mathbb{C}$–exponential tightness of the sequence of distributions of $Y^{(n)}$ and to identify a large deviation (LD) limit point as a solution to a maxingale problem, cf., Puhalskii [8]. The issue of diminishing rates is tackled by approximating trajectories that reach the boundary of the state space with trajectories that stay away from the boundary. In the process, some new techniques are developed, e.g., the LD limit point is shown to have a Markov property which enables one to identify it piecewise.

The trajectorial LDP is called upon, at first, in order to obtain an LDP for the invariant measure of $Y^{(n)}$, which is done by applying the results in Puhalskii [9]. Secondly, following the developments in Freidlin and Wentzell [6] and in Shwartz and Weiss [10], logarithmic asymptotics of both exit times from the neighbourhoods of stable equilibria and of the moments of the exit times are obtained, thus establishing metastability. As an illustration, a two–class metastable network is looked at, which is similar to the one in Antunes et al. [2].

Here is how this paper is organised. The trajectorial LDP is stated and proved in Section 2. Section 3 is concerned with the LDP for the invariant measure of $Y^{(n)}$ and the metastability. The paper uses extensively the terminology and techniques of large deviation convergence as expounded upon in Puhalskii [8]. A primer is available at the beginning of Section 3 in Puhalskii [9].

2. The trajectorial LDP

Let, for $y = (y_\theta\,, \theta \in \Theta) \in \mathbb{R}^\Theta$, $z = (z_\theta\,, \theta \in \Theta) \in \mathbb{R}^\Theta$, and $\lambda = (\lambda_\theta\,, \theta \in \Theta) \in \mathbb{R}^\Theta$,

$$\begin{aligned} H(y,\lambda) &= \sum_{k=1}^K \sum_{\theta\in\Theta_k^+} (e^{\lambda_{\theta+e_k}-\lambda_\theta}-1)\alpha_k y_\theta \\ &\quad+ \sum_{k=1}^K \sum_{\theta\in\Theta_k^-} (e^{\lambda_{\theta-e_k}-\lambda_\theta}-1)(\delta_k+\gamma_k \sum_{\theta'\in\Theta\setminus\Theta_k^+} y_{\theta'})\theta_k y_\theta \\ &\quad+ \sum_{k=1}^K \sum_{\theta'\in\Theta_k^+,\,\theta\in\Theta_k^-} (e^{\lambda_{\theta'+e_k}-\lambda_{\theta'}+\lambda_{\theta-e_k}-\lambda_\theta}-1)\theta_k\gamma_k y_\theta y_{\theta'} \end{aligned} \tag{2.1}$$

and

$$L(y,z) = \sup_{\lambda\in\mathbb{R}^\Theta} \big(\lambda\cdot z - H(y,\lambda)\big)\,, \tag{2.2}$$

where $\Theta_k^\pm = \{\theta \in \Theta : \theta \pm e_k \in \Theta\}$ and " $\cdot$ " is used to denote an inner product. Let $\mathbb{D}(\mathbb{R}_+,\mathbb{R}^\Theta)$ denote the Skorohod space of right continuous $\mathbb{R}^\Theta$–valued functions with lefthand limits. It is endowed with a metric rendering it a complete separable metric space, see, e.g., Ethier and Kurtz [5], Jacod and Shiryaev [7].

Theorem 2.1. *Let $y^{(n)} \in \mathbb{S}_{\Theta}^{(n)}$, $y \in \mathbb{S}_{|\Theta|}$, and $y^{(n)} \to y$ as $n \to \infty$. Then the sequence $Y^{(n)}$ with $Y^{(n)}(0) = y^{(n)}$ obeys an LDP in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^{\Theta})$ with deviation function*

$$I_y^*(\mathbf{y}) = \int_0^\infty L(\mathbf{y}(s), \dot{\mathbf{y}}(s))\, ds\,, \tag{2.3}$$

provided $\mathbf{y} = (\mathbf{y}(t)\,, t \geq 0)$ is an absolutely continuous function taking values in $\mathbb{S}_{|\Theta|}$ with $\mathbf{y}(0) = y$, and $I_y^*(\mathbf{y}) = \infty$, otherwise.

Remark 2.1. More explicitly, the theorem asserts that the sets $\{\mathbf{y} \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^{\Theta}) : I_y^*(\mathbf{y}) \leq \beta\}$ are compact for all $\beta \geq 0$ and that, for any Borel set $W \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R}^{\Theta})$ such that $\inf_{\mathbf{y} \in \operatorname{int} W} I_y^*(\mathbf{y}) = \inf_{\mathbf{y} \in \operatorname{cl} W} I_y^*(\mathbf{y})$,

$$(1/n) \ln \mathbf{P}(Y^{(n)} \in W) \to -\inf_{\mathbf{y} \in W} I_y^*(\mathbf{y})\,,$$

as $n \to \infty$, where int and cl denote the interior and closure of a set, respectively.

Remark 2.2. The limit of the law of large numbers in (1.1) follows with $\dot{\mathbf{y}}(t) = \nabla_\lambda H(\mathbf{y}(t), \lambda)\big|_{\lambda=0}$. In addition, $I_y^*(\mathbf{y}') = 0$ if and only if $\mathbf{y}' = \mathbf{y}$, with $\mathbf{y}(0) = y$.

A proof outline is provided next. The process $Y^{(n)}$ is a jump semimartingale. The jumps can be of several kinds: exogenous class k arrivals at nodes with population vector θ result in jumps $f_{\theta+e_k}/n - f_\theta/n$, departures of class k customers from nodes with population vector θ produce jumps $f_{\theta-e_k}/n - f_\theta/n$, whereas class k customer migrations from nodes θ to nodes θ' give rise to jumps $f_{\theta'+e_k}/n - f_{\theta'}/n + f_{\theta-e_k}/n - f_\theta/n$. Let

$$\mu^{(n)}([0,t], \Gamma) = \sum_{0<s\leq t} \mathbf{1}_{\{\Delta Y^{(n)}(s) \in \Gamma\}}$$

represent the measure of jumps of $Y^{(n)}$, where $\Delta Y^{(n)}(s) = Y^{(n)}(s) - Y^{(n)}(s-)$, with $Y^{(n)}(s-)$ denoting the lefthand limit of $Y^{(n)}$ at s and Γ standing for a Borel subset of $\mathbb{R}^{\Theta} \setminus \{0\}$. Then, assuming that $\theta \in \Theta_k^+$, $\theta \in \Theta_k^-$ and $\theta' \in \Theta_k^+$ on the lefthand sides below, where relevant,

$$\mu^{(n)}([0,t], \{\frac{f_{\theta+e_k}}{n} - \frac{f_\theta}{n}\}) = \int_0^t \sum_{i=1}^n \mathbf{1}_{\{X_i^{(n)}(s-)=\theta\}}\, dN_{i,k}^{(n)}(s)\,,$$

$$\mu^{(n)}([0,t], \{\frac{f_{\theta-e_k}}{n} - \frac{f_\theta}{n}\})$$
$$= \int_0^t \sum_{i=1}^n \mathbf{1}_{\{X_i^{(n)}(s-)=\theta\}} \sum_{j=1}^\infty \mathbf{1}_{\{j \leq X_{i,k}^{(n)}(s-)\}} dL_{i,k,j}^{(n)}(s)$$

$$+\int_0^t \sum_{\theta'\in\Theta\setminus\Theta_k^+} \sum_{i=1}^n \sum_{i'=1}^n \mathbf{1}_{\{i'\neq i\}}\mathbf{1}_{\{X_i^{(n)}(s-)=\theta\}}\mathbf{1}_{\{X_{i'}^{(n)}(s-)=\theta'\}}$$

$$\sum_{j=1}^\infty \mathbf{1}_{\{j\le X_{i,k}^{(n)}(s-)\}}\mathbf{1}_{\{\xi_{i,k,j}^{(n)}(s)=i'\}}\,dR_{i,k,j}^{(n)}(s)\,,$$

$$\mu^{(n)}([0,t],\{\frac{f_{\theta'+e_k}}{n}-\frac{f_{\theta'}}{n}+\frac{f_{\theta-e_k}}{n}-\frac{f_\theta}{n}\})$$

$$=\int_0^t \sum_{i=1}^n \sum_{i'=1}^n \mathbf{1}_{\{i'\neq i\}}\mathbf{1}_{\{X_i^{(n)}(s-)=\theta\}}\mathbf{1}_{\{X_{i'}^{(n)}(s-)=\theta'\}}$$

$$\sum_{j=1}^\infty \mathbf{1}_{\{j\le X_{i,k}^{(n)}(s-)\}}\mathbf{1}_{\{\xi_{i,k,j}^{(n)}(s)=i'\}}\,dR_{i,k,j}^{(n)}(s)\,,$$

where the $N_{i,k}^{(n)}$, $L_{i,k,j}^{(n)}$ and $R_{i,k,j}^{(n)}$ are independent Poisson processes of respective rates α_k , δ_k and γ_k , which are responsible for customer arrivals, departures and migrations, respectively, and the $\xi_{i,k,j}^{(n)}(s)$ are independent random variables uniformly distributed in $\{1,2,\dots,n\}\setminus\{i\}$, which are responsible for reroutings from node i and which are independent of the Poisson processes. The compensators of these measures of jumps relative to the natural filtration are as follows,

$$\nu^{(n)}([0,t],\{\frac{f_{\theta+e_k}}{n}-\frac{f_\theta}{n}\}) = n\int_0^t Y_\theta^{(n)}(s)\alpha_k\,ds\,,$$

$$\nu^{(n)}([0,t],\{\frac{f_{\theta-e_k}}{n}-\frac{f_\theta}{n}\})$$

$$= n\int_0^t Y_\theta^{(n)}(s)\theta_k\delta_k\,ds$$

$$+\frac{n^2}{n-1}\sum_{\theta'\in\Theta}\int_0^t Y_\theta^{(n)}(s)Y_{\theta'}^{(n)}(s)\mathbf{1}_{\{\theta'\in\Theta\setminus\Theta_k^+\}}\theta_k\gamma_k ds$$

$$-\frac{n}{n-1}\int_0^t Y_\theta^{(n)}(s)\mathbf{1}_{\{\theta\in\Theta\setminus\Theta_k^+\}}\theta_k\,\gamma_k ds\,,$$

$$\nu^{(n)}([0,t],\{\frac{f_{\theta'+e_k}}{n}-\frac{f_{\theta'}}{n}+\frac{f_{\theta-e_k}}{n}-\frac{f_\theta}{n}\})$$

$$=\frac{n^2}{n-1}\int_0^t Y_\theta^{(n)}(s)Y_{\theta'}^{(n)}(s)\theta_k\,\gamma_k\,ds$$

$$- \frac{n}{n-1} \int_0^t Y_\theta^{(n)}(s) \mathbf{1}_{\{\theta'=\theta\}} \theta_k \, \gamma_k \, ds \, .$$

Therefore, the stochastic cumulant of $Y^{(n)}$, as defined by (4.1.14) on p. 293 in Puhalskii [8], is

$$\begin{aligned}
G_t^{(n)}(\lambda) &= \int_0^t \int_{\mathbb{R}^\Theta} (e^{\lambda \cdot u} - 1) \nu^{(n)}(ds, du) \\
&= \sum_{k=1}^K \sum_{\theta \in \Theta_k^+} (e^{(\lambda_{\theta+e_k} - \lambda_\theta)/n} - 1) n \int_0^t Y_\theta^{(n)}(s) \alpha_k \, ds \\
&+ \sum_{k=1}^K \sum_{\theta \in \Theta_k^-} (e^{(\lambda_{\theta-e_k} - \lambda_\theta)/n} - 1) \big(n \int_0^t Y_\theta^{(n)}(s) \theta_k \delta_k \, ds \\
&+ \frac{n^2}{n-1} \sum_{\theta' \in \Theta} \int_0^t Y_\theta^{(n)}(s) Y_{\theta'}^{(n)}(s) \mathbf{1}_{\{\theta' \in \Theta \setminus \Theta_k^+\}} \theta_k \gamma_k ds \\
&- \frac{n}{n-1} \int_0^t Y_\theta^{(n)}(s) \mathbf{1}_{\{\theta \in \Theta \setminus \Theta_k^+\}} \theta_k \, \gamma_k ds \big) \\
&+ \sum_{k=1}^K \sum_{\theta \in \Theta_k^-, \, \theta' \in \Theta_k^+} (e^{(\lambda_{\theta'+e_k} - \lambda_{\theta'} + \lambda_{\theta-e_k} - \lambda_\theta)/n} - 1) \\
&\Big(\frac{n^2}{n-1} \int_0^t Y_\theta^{(n)}(s) Y_{\theta'}^{(n)}(s) \theta_k \, \gamma_k \, ds - \frac{n}{n-1} \int_0^t Y_\theta^{(n)}(s) \mathbf{1}_{\{\theta'=\theta\}} \theta_k \, \gamma_k \, ds \Big) \, .
\end{aligned}$$

The process $Y^{(n)}$ satisfies the hypotheses of Theorem 5.1.5 on p. 357 in Puhalskii [8]. In some more detail, since $Y^{(n)}$ is a continuous–time process, condition $(\sup \mathcal{E})$ on p. 357 in Puhalskii [8] need be checked with $\mathcal{E}_t^n(\lambda) = e^{G_t^{(n)}(\lambda)}$, see (4.1.15) on p. 293 in Puhalskii [8]. Recalling that $Y_\theta^{(n)}(t)$ takes values in $[0, 1]$ implies that the condition in question holds with

$$G_t(\lambda; \mathbf{y}) = \int_0^t H(\mathbf{y}(s), \lambda) \, ds \, .$$

By Theorem 5.1.5 on p.357 in Puhalskii [8], the sequence $Y^{(n)}$ is $\mathbb{C}$–exponentially tight in $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^\Theta)$ and its every LD limit point solves maxingale problem (y, G).

Let deviability $\Pi_y = (\Pi_y(W)\,, W \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R}^\Theta))$ represent an LD limit point of $Y^{(n)}$ (recall that $y^{(n)} \to y$), i.e., $\Pi_y(W) \in [0,1]$, $\Pi_y(\emptyset) = 0$, $\Pi_y(\mathbb{D}(\mathbb{R}_+, \mathbb{R}^\Theta)) = 1$, $\Pi_y(W) = \sup_{\mathbf{y}\in W} \Pi_y(\mathbf{y})$, sets $\{\mathbf{y} \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^\Theta) : \Pi_y(\mathbf{y}) \geq \beta\}$ are compact for $\beta \in (0,1]$, where $\Pi_y(\mathbf{y}) = \Pi_y(\{\mathbf{y}\})$, and the sequence of the distributions of $Y^{(n)}$ obeys a subsequential LDP with deviation function $-\ln \Pi_y(\mathbf{y})$, see Puhalskii [8, 9]. Then, $\Pi_y(\mathbf{y}) = 0$ unless $\mathbf{y} \in \mathbb{C}(\mathbb{R}_+, \mathbb{R}^\Theta)$, $\mathbf{y}(0) = y$, and $\exp\bigl(\lambda \cdot (\mathbf{y}(t) - y) - G_t(\lambda; \mathbf{y})\bigr)$ is a local exponential maxingale in $\mathbb{C}(\mathbb{R}_+, \mathbb{R}^\Theta)$, as defined in Puhalskii [8], where $\mathbb{C}(\mathbb{R}_+, \mathbb{R}^\Theta)$ denotes the subset of continuous functions of $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^\Theta)$ with the subspace topology which is the topology of locally uniform convergence.

By Lemma 2.7.11 on p.174 in Puhalskii [8],

$$\Pi_y(\mathbf{y}) \leq \Pi^*_y(\mathbf{y})\,, \tag{2.4}$$

where

$$\Pi^*_y(\mathbf{y}) = e^{-I^*_y(\mathbf{y})}\,. \tag{2.5}$$

It is being proved that, in fact, in (2.4) equality holds. By (2.4) and (2.5), it may be assumed that $\Pi^*_y(\mathbf{y}) > 0$. It is immediate that $\Pi^*_y(\mathbf{y}) = 0$ unless $\mathbf{y} \in \mathbb{S}_{|\Theta|}$ so that when proving the equality it may and will be assumed that $\mathbf{y}(t) = (\mathbf{y}_\theta(t)\,, \theta \in \Theta) \in \mathbb{S}_{|\Theta|}$, for all t. The equality in (2.4) is proved, at first, for the case where $\mathbf{y}$ stays away from the boundary of $\mathbb{S}_{|\Theta|}$ so that $\mathbf{y}_\theta(t) > 0$, for all θ and t, see Lemma 2.3 below. Furthermore, if $\mathbf{y}_\theta(s) > 0$ for all $s \in [0,t]$ and one defines, in analogy with pp.210, 212 in Puhalskii [8],

$$I^*_{y,t}(\mathbf{y}) = \int_0^t L(\mathbf{y}(s), \dot{\mathbf{y}}(s))\,ds\,,$$

provided $\mathbf{y}$ is absolutely continuous, $\mathbf{y}(0) = y$ and $\mathbf{y}(s) \in \mathbb{S}_{|\Theta|}$, and $I^*_{y,t}(\mathbf{y}) = \infty$, otherwise, and lets $\Pi^*_{y,t}(\mathbf{y}) = e^{-I^*_{y,t}(\mathbf{y})}$, then $\Pi_y(p_t^{-1}(p_t\mathbf{y})) = \Pi^*_{y,t}(\mathbf{y})$, where $p_t\mathbf{y} = (\mathbf{y}(s \wedge t)\,, s \geq 0)$, with $u \wedge v = \min(u, v)$. In order to tackle the case of trajectories $\mathbf{y}$ that reach the boundary of the state space, one needs to find trajectories $\mathbf{y}^\varepsilon$ that are locally bounded away from zero entrywise and converge to $\mathbf{y}$ locally uniformly, as $\varepsilon \to 0$, such that

$$\int_0^t L(\mathbf{y}^\varepsilon(s), \dot{\mathbf{y}}^\varepsilon(s))\,ds \to \int_0^t L(\mathbf{y}(s), \dot{\mathbf{y}}(s))\,ds\,. \tag{2.6}$$

The hard part in the proof of (2.6) is verifying the hypotheses of Lebesgue's dominated convergence theorem. The needed majoration for the $L(\mathbf{y}^\varepsilon(s), \dot{\mathbf{y}}^\varepsilon(s))$ is obtained through the use of a nontrivial bound on the optimisers in (2.2), see Lemma 2.2. Upper semicontinuity of $\Pi_y(\mathbf{y})$ in $(y, \mathbf{y})$ is also important and novel, see Lemma 2.6. Nevertheless, even then the convergence in (2.6) is proved for values of t that are not too great. In order to finish the proof of Theorem 2.1, an arbitrary trajectory is cut into pieces, for each of which $\Pi_y(p_t^{-1}(p_t\mathbf{y})) = \Pi^*_{y,t}(\mathbf{y})$, and a Markov property of Π_y is used in order to obtain the needed equality $\Pi_y(\mathbf{y})) = \Pi^*_y(\mathbf{y})$.

Lemma 2.1. *Let $y \in \mathbb{S}_{|\Theta|}$ and $z \in \mathbb{R}^\Theta$. If $\sum_{\theta\in\Theta} z_\theta = 0$ and $y_\theta > 0$, for all θ, then supremum in* (2.2) *is attained.*

Proof. Since $\sum_{\theta\in\Theta} z_\theta = 0$, it may be assumed that $\lambda_0 = 0$. If $|\lambda_\theta| \to \infty$, for certain θ, and $|\lambda_{\hat\theta}|$ grows the fastest, then there exist $\tilde\theta$ and k such that either $|\lambda_{\tilde\theta+e_k} - \lambda_{\tilde\theta}|$ or $|\lambda_{\tilde\theta-e_k} - \lambda_{\tilde\theta}|$ tends to infinity at the same rate or faster, which implies that either $|\lambda_{\hat\theta} z_{\hat\theta}| - e^{\lambda_{\tilde\theta+e_k}-\lambda_{\tilde\theta}}\alpha_k y_{\tilde\theta} - e^{\lambda_{\tilde\theta}-\lambda_{\tilde\theta+e_k}}\delta_k y_{\tilde\theta+e_k} \to -\infty$ or $|\lambda_{\hat\theta} z_{\hat\theta}| - e^{\lambda_{\tilde\theta-e_k}-\lambda_{\tilde\theta}}\delta_k y_{\tilde\theta} - e^{\lambda_{\tilde\theta}-\lambda_{\tilde\theta-e_k}}\alpha_k y_{\tilde\theta-e_k} \to -\infty$. Hence, supremum in (2.2) may be taken over a bounded set, so, it is attained. □

Remark 2.3. One can prove the following criterion for $L(y,z)$ to be finite. Given $\theta,\theta' \in \Theta$, a sequence $\theta_0,\theta_1,\dots,\theta_\ell$ of elements of Θ is called a path from θ to θ' provided $\theta_0 = \theta$, $\theta_\ell = \theta'$ and either $\theta_{i+1} = \theta_i + e_{k_i}$ or $\theta_{i+1} = \theta_i - e_{k_i}$, for some k_i, for all $i \in \{0,1,2,\dots,\ell-1\}$. For $y \in \mathbb{S}_{|\Theta|}$, it is said that points θ and θ' y–communicate if $y_\theta > 0$, $y_{\theta'} > 0$ and there exists a path from θ to θ' such that $y_{\tilde\theta} > 0$, for every $\tilde\theta$ on the path. The communication relation is an equivalence relation. The equivalence classes are denoted by $\Theta_1(y),\dots,\Theta_{m(y)}(y)$. Suppose that $z_\theta = 0$ when $y_\theta = 0$. Then $L(y,z) < \infty$ if and only if $\sum_{\theta\in\Theta_i(y)} z_\theta = 0$, for each $i = 1,2,\dots,m(y)$.

Lemma 2.2. *There exist C_1 and C_2 such that if λ delivers supremum in* (2.2), *then for all k, $\theta \in \Theta_k^+$, and $\theta' \in \Theta_k^-$,*

$$e^{\lambda_{\theta+e_k}-\lambda_\theta} y_\theta + e^{\lambda_{\theta'-e_k}-\lambda_{\theta'}} y_{\theta'} + e^{\lambda_{\theta+e_k}-\lambda_\theta+\lambda_{\theta'-e_k}-\lambda_{\theta'}} y_\theta y_{\theta'} \le C_1 + C_2 \sum_{\theta''\in\Theta} |z_{\theta''}|.$$

Proof. For $\hat\lambda = (\hat\lambda_{\theta,k}) \in \prod_{k\in\{1,2,\dots,K\}} \mathbb{R}^{\Theta_k^+}$, define

$$\begin{aligned}\hat H(y,\hat\lambda) = &\sum_{k=1}^K \sum_{\theta\in\Theta_k^+} (e^{\hat\lambda_{\theta,k}} - 1)\alpha_k y_\theta \\ &+ \sum_{k=1}^K \sum_{\theta\in\Theta_k^-} (e^{-\hat\lambda_{\theta-e_k,k}} - 1)(\delta_k + \gamma_k \sum_{\theta'\in\Theta\setminus\Theta_k^+} y_{\theta'})\theta_k y_\theta \\ &+ \sum_{k=1}^K \sum_{\theta\in\Theta_k^-,\,\theta'\in\Theta_k^+} (e^{\hat\lambda_{\theta',k}-\hat\lambda_{\theta-e_k,k}} - 1)\theta_k\gamma_k y_\theta y_{\theta'}\end{aligned}$$

so that

$$L(y,z)=\sup_{\lambda=(\lambda_\theta)\in\mathbb{R}^\Theta}\bigl(\lambda\cdot z-H(y,\lambda)\bigr)=\sup_{\substack{\lambda=(\lambda_\theta)\in\mathbb{R}^\Theta\,,\\ \hat\lambda=(\hat\lambda_{\theta,k})\in\prod_{k\in\{1,2,\dots,K\}}\mathbb{R}^{\Theta_k^+}:\\ \lambda_{\theta+e_k}-\lambda_\theta-\hat\lambda_{\theta,k}=0}}\bigl(\lambda\cdot z-\hat H(y,\hat\lambda)\bigr)\,. \tag{2.7}$$

Define a Lagrange function, with $r_{\theta,k}\in\mathbb{R}$ and $r=(r_{\theta,k})$,

$$\mathcal{L}(\lambda,\hat\lambda,r,y,z)=\sum_{\theta\in\Theta}\lambda_\theta z_\theta-\hat H(y,\hat\lambda)+\sum_{\theta\in\Theta_k^+}r_{\theta,k}(\lambda_{\theta+e_k}-\lambda_\theta-\hat\lambda_{\theta,k})\,.$$

The optimality conditions in (2.7) that $\partial_{\hat\lambda_{\theta,k}}\mathcal{L}(\lambda,\hat\lambda,r,y,z)=0$ and $\partial_{\lambda_\theta}\mathcal{L}(\lambda,\hat\lambda, r,y,z)=0$, see, e.g., Theorem 3.2.2 on p.253 in Alekseev et al. [1], imply that, for $\theta\in\Theta_k^+$, there exist $r_{\theta,k}$ such that

$$\begin{aligned}&-e^{\hat\lambda_{\theta,k}}(\alpha_k+\gamma_k\sum_{\theta'\in\Theta_k^+}e^{-\hat\lambda_{\theta',k}}(\theta'_k+1)y_{\theta'+e_k})y_\theta\\&\quad+e^{-\hat\lambda_{\theta,k}}(\theta_k+1)\bigl(\delta_k+\gamma_k\sum_{\theta'\in\Theta\setminus\Theta_k^+}y_{\theta'}+\gamma_k\sum_{\theta'\in\Theta_k^+}e^{\hat\lambda_{\theta',k}}y_{\theta'}\bigr)y_{\theta+e_k}=r_{\theta,k}\end{aligned} \tag{2.8}$$

and

$$z_\theta+r_{\theta-e_k,k}-r_{\theta,k}=0\,,\text{ for }\theta\in\Theta_k^+\text{ with }\theta_k\ge1\,, \tag{2.9a}$$

$$z_\theta-r_{\theta,k}=0\,,\text{ for }\theta\in\Theta_k^+\text{ with }\theta_k=0\,. \tag{2.9b}$$

Summing in (2.8) yields

$$-\sum_{\theta\in\Theta_k^+}e^{\hat\lambda_{\theta,k}}\alpha_k y_\theta+\sum_{\theta\in\Theta_k^+}e^{-\hat\lambda_{\theta,k}}(\delta_k+\gamma_k\sum_{\theta'\in\Theta\setminus\Theta_k^+}y_{\theta'})(\theta_k+1)y_{\theta+e_k}=\sum_{\theta\in\Theta}r_{\theta,k}\,. \tag{2.10}$$

Solving for $\sum_{\theta\in\Theta_k^+}e^{\hat\lambda_{\theta,k}}y_\theta$ and substituting in (2.8) imply, after some algebra, that, for $\theta\in\Theta_k^+$,

$$\begin{aligned}&\alpha_k e^{\hat\lambda_{\theta,k}}y_\theta\\&\quad-e^{-\hat\lambda_{\theta,k}}y_{\theta+e_k}(\theta_k+1)\\&\qquad\times\bigl(\delta_k+\gamma_k\sum_{\theta'\in\Theta\setminus\Theta_k^+}y_{\theta'}-\frac{\gamma_k\sum_{\theta'}r_{\theta',k}}{\alpha_k+\gamma_k\sum_{\theta'\in\Theta_k^+}e^{-\hat\lambda_{\theta',k}}(\theta'_k+1)y_{\theta'+e_k}}\bigr)\\&\qquad=\frac{-\alpha_k r_{\theta,k}}{\alpha_k+\gamma_k\sum_{\theta'\in\Theta_k^+}e^{-\hat\lambda_{\theta',k}}(\theta'_k+1)y_{\theta'+e_k}}\,.\end{aligned}$$

Therefore, if $\hat{\lambda}_{\theta,k} > 0$, then

$$\begin{aligned}\alpha_k e^{\hat{\lambda}_{\theta,k}} y_\theta \le{}& y_{\theta+e_k}(\theta_k+1)\\ &\times \big(\delta_k + \gamma_k \sum_{\theta'\in\Theta\setminus\Theta_k^+} y_{\theta'} + \frac{\gamma_k|\sum_{\theta'} r_{\theta',k}|}{\alpha_k + \gamma_k\sum_{\theta'\in\Theta_k^+} e^{-\hat{\lambda}_{\theta',k}}(\theta'_k+1)y_{\theta'+e_k}}\big)\\ &- \frac{\alpha_k r_{\theta,k}}{\alpha_k + \gamma_k\sum_{\theta'\in\Theta_k^+} e^{-\hat{\lambda}_{\theta',k}}(\theta'_k+1)y_{\theta'+e_k}}\\ &\le y_{\theta+e_k}(\theta_k+1)\big(\delta_k+\gamma_k+\frac{\gamma_k|\sum_{\theta'} r_{\theta',k}|}{\alpha_k}\big) + |r_{\theta,k}|\,,\end{aligned}$$

which implies that, no matter the sign of $\hat{\lambda}_{\theta,k}$,

$$\alpha_k e^{\hat{\lambda}_{\theta,k}} y_\theta \le \alpha_k y_\theta + y_{\theta+e_k}(\theta_k+1)\big(\delta_k+\gamma_k+\frac{\gamma_k|\sum_{\theta'} r_{\theta',k}|}{\alpha_k}\big) + |r_{\theta,k}|\,. \tag{2.11}$$

By (2.10) and (2.11),

$$\begin{aligned}&\sum_{\theta\in\Theta} e^{-\hat{\lambda}_{\theta,k}}\delta_k(\theta_k+1)y_{\theta+e_k}\\ &\quad\le \sum_{\theta\in\Theta}\big(\alpha_k y_\theta + y_{\theta+e_k}(\theta_k+1)\big(\delta_k+\gamma_k+\frac{\gamma_k|\sum_{\theta'} r_{\theta',k}|}{\alpha_k}\big)+|r_{\theta,k}|\big)\\ &\qquad+\sum_{\theta\in\Theta} r_{\theta,k}\,.\end{aligned} \tag{2.12}$$

Solving (2.9a) and (2.9b) recursively yields

$$r_{\theta,k} = \sum_{i=0}^{\theta_k} z_{\theta-ie_k}\,. \tag{2.13}$$

As a consequence of (2.11), (2.12), and (2.13), for some $C'_1 > 0$ and $C'_2 > 0$,

$$e^{\hat{\lambda}_{\theta,k}} y_\theta + e^{-\hat{\lambda}_{\theta,k}} y_{\theta+e_k} \le C'_1 + C'_2\sum_{\theta'\in\Theta}|z_{\theta'}|\,. \tag{2.14}$$

By (2.8), in analogy with (2.11), for $\theta\in\Theta_k^+$,

$$\begin{aligned}&e^{\hat{\lambda}_{\theta,k}}\big(\alpha_k+\gamma_k\sum_{\theta'\in\Theta_k^+} e^{-\hat{\lambda}_{\theta',k}}(\theta'_k+1)y_{\theta'+e_k}\big)y_\theta\\ &\quad\le(\theta_k+1)\big(\delta_k+\gamma_k\sum_{\theta'\in\Theta\setminus\Theta_k^+} y_{\theta'}+\gamma_k\sum_{\theta'\in\Theta_k^+} e^{\hat{\lambda}_{\theta',k}}y_{\theta'}\big)y_{\theta+e_k}+|r_{\theta,k}|\end{aligned}$$

$$+ (\alpha_k + \gamma_k \sum_{\theta' \in \Theta_k^+} e^{-\hat{\lambda}_{\theta',k}}(\theta'_k + 1)y_{\theta'+e_k})y_\theta \,.$$

By (2.13) and (2.14), there exist $C''_1 > 0$ and $C''_2 > 0$ such that, for $\theta \in \Theta_k^+$ and $\theta' \in \Theta_k^-$,

$$e^{\hat{\lambda}_{\theta,k} - \hat{\lambda}_{\theta'-e_k,k}} y_\theta y_{\theta'} \le C''_1 + C''_2 \sum_{\theta'' \in \Theta} |z_{\theta''}| \,,$$

which concludes the proof on recalling that $\hat{\lambda}_{\theta,k} = \lambda_{\theta+e_k} - \lambda_\theta$ and that $\hat{\lambda}_{\theta'-e_k,k} = \lambda_{\theta'} - \lambda_{\theta'-e_k}$. □

Denote $\Pi_{y,t}(\mathbf{y}) = \Pi_y(p_t^{-1}(p_t \mathbf{y}))$.

Lemma 2.3. *Let $\mathbf{y} = (\mathbf{y}(t)\,, t \ge 0)$ be an absolutely continuous function taking values in $\mathbb{S}_{|\Theta|}$ with $\mathbf{y}(0) = y$. If $\mathbf{y}$ is locally bounded away from zero entrywise , then $\Pi_{y,t}(\mathbf{y}) = \Pi^*_{y,t}(\mathbf{y})$, for all t , and $\Pi_y(\mathbf{y}) = \Pi^*_y(\mathbf{y})$.*

Proof. By Lemma 2.1, there exists function $(\lambda(s)\,, s \ge 0) = ((\lambda_\theta(s), \theta \in \Theta),\ s \ge 0)$ such that, a.e.,

$$L(\mathbf{y}(s), \dot{\mathbf{y}}(s)) = \sum_{\theta \in \Theta} \lambda_\theta(s) \dot{\mathbf{y}}_\theta(s) - H(\mathbf{y}(s), \lambda(s)) \,,$$

with "almost everywhere" here and below being understood with respect to the Lebesgue measure. The following equation is satisfied a.e.:

$$\dot{\mathbf{y}}(s) = \nabla_\lambda H(\mathbf{y}(s), \lambda(s)) \,. \tag{2.15}$$

Calculations, using (2.1), yield

$$\begin{aligned}
\dot{\mathbf{y}}_\theta(s) = (\alpha_k &+ \sum_{\theta' \in \Theta_k^-} \theta'_k \gamma_k e^{\lambda_{\theta'-e_k}(s) - \lambda_{\theta'}(s)} \mathbf{y}_{\theta'}(s)) e^{\lambda_\theta(s) - \lambda_{\theta-e_k}(s)} \mathbf{y}_{\theta-e_k}(s) \mathbf{1}_{\{\theta \in \Theta_k^-\}} \\
&+ \big(\delta_k + \gamma_k \sum_{\theta \in \Theta \setminus \Theta_k^+} \mathbf{y}_{\theta'}(s) + \gamma_k \sum_{\theta' \in \Theta_k^+} e^{\lambda_{\theta'+e_k}(s) - \lambda_{\theta'}(s)} \mathbf{y}_{\theta'}(s)\big) \\
&\quad \times e^{\lambda_\theta(s) - \lambda_{\theta+e_k}(s)} (\theta_k + 1) \mathbf{y}_{\theta+e_k}(s) \mathbf{1}_{\{\theta \in \Theta_k^+\}} \\
&- \big((\alpha_k + \sum_{\theta' \in \Theta_k^-} \theta'_k \gamma_k e^{\lambda_{\theta'-e_k}(s) - \lambda_{\theta'}(s)} \mathbf{y}_{\theta'}(s)) e^{\lambda_{\theta+e_k}(s) - \lambda_\theta(s)} \mathbf{1}_{\{\theta \in \Theta_k^+\}} \\
&+ (\delta_k + \gamma_k \sum_{\theta' \in \Theta \setminus \Theta_k^+} \mathbf{y}_{\theta'}(s) + \gamma_k \sum_{\theta' \in \Theta_k^+} \\
&\quad \times e^{\lambda_{\theta'+e_k}(s) - \lambda_{\theta'}(s)} \mathbf{y}_{\theta'}(s)) e^{\lambda_{\theta-e_k}(s) - \lambda_\theta(s)} \theta_k \mathbf{1}_{\{\theta \in \Theta_k^-\}} \big) \mathbf{y}_\theta(s) \,.
\end{aligned}$$

Since the $\mathbf{y}_\theta(s)$ are locally bounded away from zero, Lemma 2.2 implies that the exponentials on the latter righthand side are locally integrable functions of

s, so, the righthand side of (2.15) is a Lipschitz continuous function of $\mathbf{y}(s)$. It follows that $\mathbf{y}$ is a unique solution of (2.15). By Theorem 2.8.14 on p.213 and Lemma 2.8.20 on p.218 in Puhalskii [8], $\Pi_{y,t}(\mathbf{y}) = \Pi^*_{y,t}(\mathbf{y})$ and $\Pi_y(\mathbf{y}) = \Pi^*_y(\mathbf{y})$. □

Lemma 2.4. *Let* $\mathbf{y}$ *be an absolutely continuous function with values in* $\mathbb{S}_{|\Theta|}$. *Let* θ^* *represent a point of the maximum of* $\mathbf{y}_\theta(0)$ *so that* $\mathbf{y}_{\theta^*}(0) = \max_{\theta\in\Theta}\mathbf{y}_\theta(0)$. *For* $\varepsilon \in (0, 1/(3|\Theta|^2))$, *let* $\mathbf{y}^\varepsilon_\theta(s) = \mathbf{y}_\theta(s) + \varepsilon$ *unless* $\theta = \theta^*$ *and let* $\mathbf{y}^\varepsilon_{\theta^*}(s) = 1 - \sum_{\theta\neq\theta^*}\mathbf{y}^\varepsilon_\theta(s)$. *Then, for* t *such that*

$$\sum_{\theta\in\Theta}\int_0^t |\dot{\mathbf{y}}_\theta(s)|\, ds \le 1/(3|\Theta|)\,,$$

$\mathbf{y}^\varepsilon(s) \in \mathbb{S}_{|\Theta|}$ *on* $[0,t]$ *and*

$$\lim_{\varepsilon\to 0}\int_0^t L(\mathbf{y}^\varepsilon(s), \dot{\mathbf{y}}^\varepsilon(s))\, ds = \int_0^t L(\mathbf{y}(s), \dot{\mathbf{y}}(s))\, ds\,. \tag{2.16}$$

Proof. The functions $\mathbf{y}^\varepsilon_{\theta^*}(s)$ are bounded away from zero on $[0,t]$, uniformly over ε. Indeed, since $\mathbf{y}_{\theta^*}(0) \ge 1/|\Theta|$, $\mathbf{y}_{\theta^*}(s) \ge 2/(3|\Theta|)$ on $[0,t]$. It follows that $\mathbf{y}^\varepsilon_{\theta^*}(s) \ge 2/(3|\Theta|) - \varepsilon|\Theta| \ge 1/(3|\Theta|)$ on $[0,t]$. Evidently, $\dot{\mathbf{y}}^\varepsilon_\theta(s) = \dot{\mathbf{y}}_\theta(s)$ a.e. and $\mathbf{y}^\varepsilon_\theta(s) \to \mathbf{y}_\theta(s)$ uniformly on bounded sets, for all θ, as $\varepsilon \to 0$. By Lemma 2.1, a.e., the supremum in (2.2) with $y = \mathbf{y}^\varepsilon(s)$ and $z = \dot{\mathbf{y}}^\varepsilon(s)$ is attained at some $\lambda^\varepsilon(s)$. Since $\mathbf{y}^\varepsilon_\theta(s) = \mathbf{y}_\theta(s) + \varepsilon$, for $\theta \neq \theta^*$, and $\mathbf{y}^\varepsilon_{\theta^*}(s) = \mathbf{y}_{\theta^*}(s) - (|\Theta| - 1)\varepsilon$, the definition of $H(y,\lambda)$ in (2.1) implies that

$$\begin{aligned}
H(\mathbf{y}^\varepsilon(s), \lambda) \ge & \sum_{k=1}^K\Big(\sum_{\theta\in\Theta_k^+} e^{\lambda_{\theta+e_k}-\lambda_\theta}\alpha_k\mathbf{y}_\theta(s) - |\Theta|\varepsilon e^{\lambda_{\theta^*+e_k}-\lambda_{\theta^*}}\alpha_k\Big)\\
& - \sum_{k=1}^K\sum_{\theta\in\Theta_k^+}\alpha_k(\mathbf{y}_\theta(s)+\varepsilon)\\
& + \sum_{k=1}^K\Big(\sum_{\theta\in\Theta_k^-} e^{\lambda_{\theta-e_k}-\lambda_\theta}(\delta_k + \gamma_k(\sum_{\theta'\in\Theta\setminus\Theta_k^+}\mathbf{y}_{\theta'}(s) - \varepsilon|\Theta|))\theta_k\mathbf{y}_\theta(s)\\
& - e^{\lambda_{\theta^*-e_k}-\lambda_{\theta^*}}(\delta_k+\gamma_k)\theta^*_k|\Theta|\varepsilon\Big)\\
& - \sum_{k=1}^K\sum_{\theta\in\Theta_k^-}(\delta_k + \gamma_k\sum_{\theta'\in\Theta\setminus\Theta_k^+}\mathbf{y}_{\theta'}(s) + \gamma_k\varepsilon|\Theta|)\theta_k(\mathbf{y}_\theta(s)+\varepsilon)\\
& + \sum_{k=1}^K\gamma_k\big(\sum_{\theta'\in\Theta_k^+} e^{\lambda_{\theta'+e_k}-\lambda_{\theta'}}\mathbf{y}_{\theta'}(s) - e^{\lambda_{\theta^*+e_k}-\lambda_{\theta^*}}|\Theta|\varepsilon\big)
\end{aligned}$$

$$\times \Big(\sum_{\theta \in \Theta_k^-} e^{\lambda_{\theta - e_k} - \lambda_\theta} \theta_k \mathbf{y}_\theta(s) - e^{\lambda_{\theta^* - e_k} - \lambda_{\theta^*}} \theta_k^* |\Theta| \varepsilon \Big)$$
$$- \sum_{k=1}^K \sum_{\theta \in \Theta_k^-, \theta' \in \Theta_k^+} \theta_k \gamma_k (\mathbf{y}_\theta(s) + \varepsilon)(\mathbf{y}_{\theta'}(s) + \varepsilon)$$
$$\geq H(\mathbf{y}(s), \lambda) - |\Theta| \varepsilon\, R(\mathbf{y}(s), \lambda) - \varepsilon M\,,$$

where

$$R(y, \lambda) = \sum_{k=1}^K e^{\lambda_{\theta^* + e_k} - \lambda_{\theta^*}} \alpha_k + \sum_{k=1}^K \sum_{\theta \in \Theta_k^-} e^{\lambda_{\theta - e_k} - \lambda_\theta} \theta_k y_\theta$$
$$+ \sum_{k=1}^K e^{\lambda_{\theta^* - e_k} - \lambda_{\theta^*}} (\delta_k + \gamma_k) \theta_k^*$$
$$+ \sum_{k=1}^K \gamma_k e^{\lambda_{\theta^* + e_k} - \lambda_{\theta^*}} \sum_{\theta \in \Theta_k^-} e^{\lambda_{\theta - e_k} - \lambda_\theta} \theta_k y_\theta(s)$$
$$+ \sum_{k=1}^K \gamma_k \sum_{\theta \in \Theta_k^+} e^{\lambda_{\theta + e_k} - \lambda_\theta} y_\theta(s) e^{\lambda_{\theta^* - e_k} - \lambda_{\theta^*}} \theta_k^*$$

and $M > 0$ depends neither on λ nor on $\mathbf{y}(s)$.

Therefore,

$$L(\mathbf{y}^\varepsilon(s), \dot{\mathbf{y}}^\varepsilon(s)) = \sum_{\theta \in \Theta} \lambda_\theta^\varepsilon(s) \dot{\mathbf{y}}_\theta^\varepsilon(s) - H(\mathbf{y}^\varepsilon(s), \lambda^\varepsilon(s))$$
$$\leq \sum_{\theta \in \Theta} \lambda_\theta^\varepsilon(s) \dot{\mathbf{y}}_\theta(s) - H(\mathbf{y}(s), \lambda^\varepsilon(s))$$
$$+ |\Theta| \varepsilon\, R(\mathbf{y}(s), \lambda^\varepsilon(s)) + \varepsilon M$$
$$\leq L(\mathbf{y}(s), \dot{\mathbf{y}}(s)) + |\Theta| \varepsilon\, R(\mathbf{y}(s), \lambda^\varepsilon(s)) + \varepsilon M\,.$$

Since $\mathbf{y}_{\theta^*}^\varepsilon(s)$ is locally bounded away from zero on $[0, t]$ uniformly in ε, Lemma 2.2 implies that

$$\limsup_{\varepsilon \to 0} \int_0^t R(\mathbf{y}(s), \lambda^\varepsilon(s))\, ds < \infty\,.$$

Thus,

$$\limsup_{\varepsilon \to 0} \int_0^t L(\mathbf{y}^\varepsilon(s), \dot{\mathbf{y}}^\varepsilon(s))\, ds \leq \int_0^t L(\mathbf{y}(s), \dot{\mathbf{y}}(s))\, ds\,. \tag{2.17}$$

On the other hand, by (2.1) and (2.2), a.e.,

$$\liminf_{\varepsilon\to 0} L(\mathbf{y}^\varepsilon(s), \dot{\mathbf{y}}^\varepsilon(s)) \geq L(\mathbf{y}(s), \dot{\mathbf{y}}(s))\,.$$

When put together with (2.17) and Fatou's lemma, this proves (2.16). □

The function $\mathbf{y}^\varepsilon$ in the above lemma can be used as an approximation for $\mathbf{y}$ until $\mathbf{y}_{\theta^*}(s)$ hits 0. At that stage, one starts afresh by choosing different θ as θ^*. The piecing together is done with the use of the Markov property in the following lemma.

Lemma 2.5. $(\Pi_y\,, y\in\mathbb{S}_{|\Theta|})$ *is an idempotent Markov family in the sense that, for* $\mathbf{y}\in\mathbb{D}(\mathbb{R}_+,\mathbb{S}_{|\Theta|})$,

$$\Pi_y(\mathbf{y}) = \Pi_{y,t}(\mathbf{y})\Pi_{\mathbf{y}_t}(\vartheta_t\mathbf{y})\,,$$

where $\vartheta_t\mathbf{y} = (\mathbf{y}(s+t)\,, s\geq 0)$.

Proof. Let $f(\mathbf{y})$ represent a nonnegative, bounded and continuous function on $\mathbb{D}(\mathbb{R}_+,\mathbb{R}^\Theta)$. By the Markov property of $Y^{(n)}$, on writing $\mathbf{y} = (p_t\mathbf{y}, \vartheta_t\mathbf{y})$, with $E_{y^{(n)}}$ representing expectation when $Y^{(n)}$ starts at $y^{(n)}$ and $(\mathcal{F}^{(n)}(t)\,, t\geq 0)$ representing the filtration associated with $Y^{(n)}$, provided $y^{(n)}\in\mathbb{S}^{(n)}_{|\Theta|}$,

$$\begin{aligned} E_{y^{(n)}}(f(Y^{(n)}))^n &= E_{y^{(n)}}E_{y^{(n)}}((f((p_tY^{(n)}, \vartheta_tY^{(n)})))^n|\mathcal{F}^{(n)}(t)) \\ &= E_{y^{(n)}}\big(E_{Y^{(n)}(t)}f(u, Y^{(n)})^n\Big|_{u=p_tY^{(n)}}\big)\,. \end{aligned} \tag{2.18}$$

By Theorem 2.1, if $y^{(n)}\to y$ in $\mathbb{R}^\Theta$ and $u_n\to u$ in $\mathbb{D}(\mathbb{R}_+,\mathbb{R}^\Theta)$, as $n\to\infty$, then

$$(E_{y^{(n)}}f(p_tu_n, Y^{(n)})^n)^{1/n} \to \sup_{\tilde{\mathbf{y}}} f(p_tu, \tilde{\mathbf{y}})\Pi_y(\tilde{\mathbf{y}})\,.$$

With $g_n(\mathbf{y}) = (E_{\mathbf{y}(t)}f(p_t\mathbf{y}, Y^{(n)})^n)^{1/n}$, if $\mathbf{y}^{(n)}\to\mathbf{y}$, then

$$g_n(\mathbf{y}^{(n)}) \to \sup_{\tilde{\mathbf{y}}} f(p_t\mathbf{y}, \tilde{\mathbf{y}})\Pi_{\mathbf{y}(t)}(\tilde{\mathbf{y}})\,.$$

Therefore, accounting for (2.18),

$$\begin{aligned} \big(E_{y^{(n)}}f(Y^{(n)})^n\big)^{1/n} &= \big(E_{y^{(n)}}(g_n(Y^{(n)}))^n\big)^{1/n} \\ &\to \sup_{\mathbf{y}}\sup_{\tilde{\mathbf{y}}} f(p_t\mathbf{y}, \tilde{\mathbf{y}})\Pi_{\mathbf{y}(t)}(\tilde{\mathbf{y}})\Pi_y(\mathbf{y}) \\ &= \sup_{\mathbf{y}}\sup_{\tilde{\mathbf{y}}} f(p_t\mathbf{y}, \tilde{\mathbf{y}})\Pi_{\mathbf{y}(t)}(\tilde{\mathbf{y}})\Pi_y(p_t^{-1}(p_t\mathbf{y})) \\ &= \sup_{\mathbf{y},\mathbf{y}'} f(p_t\mathbf{y}, \vartheta_t\mathbf{y}')\Pi_{\mathbf{y}(t)}(\vartheta_t\mathbf{y}')\Pi_y(p_t^{-1}(p_t\mathbf{y})) \end{aligned}$$

$$= \sup_{\mathbf{y}} f(p_t\mathbf{y}, \vartheta_t\mathbf{y})\Pi_{\mathbf{y}(t)}(\vartheta_t\mathbf{y})\Pi_y(p_t^{-1}(p_t\mathbf{y}))$$
$$= \sup_{\mathbf{y}} f(\mathbf{y})\Pi_{\mathbf{y}(t)}(\vartheta_t\mathbf{y})\Pi_y(p_t^{-1}(p_t\mathbf{y}))\,.$$

□

Lemma 2.6. *The function* $\Pi_y(\mathbf{y})$ *is upper semicontinuous in* $(y,\mathbf{y})$.

Proof. Suppose that initial conditions $Y^{(n)}(0)$ are independent of the random entities driving the processes $Y^{(n)}$ and satisfy an LDP in $\mathbb{R}^\Theta$ with a continuous deviation function I^Y . Then the distributions of the pairs $(Y^{(n)}(0), Y^n)$ satisfy a subsequential LDP with $I^Y(y) - \ln \Pi_y(\mathbf{y})$. Since the latter quantity is lower semicontinuous in $(y,\mathbf{y})$ and $I^Y(y)$ is continuous in y , $\Pi_y(\mathbf{y})$ is upper semicontinuous in $(y,\mathbf{y})$. □

Proof of Theorem 2.1. Since $\Pi_y(\mathbf{y}) = \lim_{t\to\infty} \Pi_{y,t}(\mathbf{y})$ and $\Pi^*_y(\mathbf{y}) = \lim_{t\to\infty} \Pi^*_{y,t}(\mathbf{y})$, it suffices to prove that

$$\Pi_{y,t}(\mathbf{y}) = \Pi^*_{y,t}(\mathbf{y})\,. \tag{2.19}$$

It is shown, first, that (2.19) holds for all t such that $\int_0^t \sum_{\theta\in\Theta} |\dot{\mathbf{y}}_\theta(s)|\,ds \le 1/(3|\Theta|)$. By Lemma 2.4, there exist $\mathbf{y}^\varepsilon$, such that $\mathbf{y}^\varepsilon_\theta(s) > 0$, for all θ , on $[0,t]$, $\mathbf{y}^\varepsilon(s) \to \mathbf{y}(s)$ on $[0,t]$ and $\Pi^*_{\mathbf{y}^\varepsilon(0),t}(\mathbf{y}^\varepsilon) \to \Pi^*_{y,t}(\mathbf{y})$, as $\varepsilon \to 0$. Since, by Lemma 2.3, $\Pi_{\mathbf{y}^\varepsilon(0)}(p_t^{-1}(p_t\mathbf{y}^\varepsilon)) = \Pi^*_{\mathbf{y}^\varepsilon(0),t}(\mathbf{y}^\varepsilon)$, by upper semicontinuity, $\Pi_y(p_t^{-1}(p_t\mathbf{y})) \ge \limsup_{\varepsilon\to 0} \Pi_{\mathbf{y}^\varepsilon(0)}(p_t^{-1}(p_t\mathbf{y}^\varepsilon)) = \limsup_{\varepsilon\to0} \Pi^*_{\mathbf{y}^\varepsilon(0),t}(\mathbf{y}^\varepsilon) = \Pi^*_{y,t}(\mathbf{y})$. On the other hand, $\Pi_y(p_t^{-1}(p_t\mathbf{y})) \le \Pi^*_{y,t}(\mathbf{y})$, generally, proving (2.19).

Given arbitrary $t > 0$, there exist $0 = t_0 < t_1 < \ldots < t_m = t$ such that $\int_{t_{i-1}}^{t_i} \sum_{\theta\in\Theta} |\dot{\mathbf{y}}_\theta(s)|\,ds \le 1/(3|\Theta|)$, for all $i = 1,\ldots,m$. Consequently, by the argument in the preceding paragraph, $\Pi_{\mathbf{y}(t_{i-1})}(p^{-1}_{t_i-t_{i-1}}(p_{t_i-t_{i-1}}(\vartheta_{t_{i-1}}\mathbf{y}))) = \Pi^*_{\mathbf{y}(t_{i-1}),t_i-t_{i-1}}(\vartheta_{t_{i-1}}\mathbf{y})$. By Lemma 2.5,

$$\Pi_y(p_t^{-1}(p_t\mathbf{y})) = \prod_{i=1}^m \Pi_{\mathbf{y}(t_{i-1})}(p^{-1}_{t_i-t_{i-1}}(p_{t_i-t_{i-1}}(\vartheta_{t_{i-1}}\mathbf{y})))$$
$$= \prod_{i=1}^m \Pi^*_{\mathbf{y}(t_{i-1}),t_i-t_{i-1}}(\vartheta_{t_{i-1}}\mathbf{y}) = \Pi^*_{y,t}(\mathbf{y})\,.$$

□

3. Large deviations of the invariant measure. Metastability

Being irreducible and having a finite state space, the process $Y^{(n)}$ possesses a unique invariant measure on $\mathbb{S}^{(n)}_{|\Theta|}$, see, e.g., Asmussen [3], which is denoted by $\mu^{(n)}$. It is convenient to extend $\mu^{(n)}$ to the whole of $\mathbb{S}_{|\Theta|}$ by letting $\mu^{(n)}(\mathbb{S}_{|\Theta|} \setminus$

$\mathbb{S}^{(n)}_{|\Theta|}) = 0$. The results in Puhalskii [9] enable one to obtain large deviation asymptotics of $\mu^{(n)}$.

Let, for $t > 0$, $y \in \mathbb{S}_{|\Theta|}$ and $y' \in \mathbb{S}_{|\Theta|}$,

$$\Phi_t(y, y') = \inf_{\substack{\mathbf{y}\in\mathbb{C}(\mathbb{R}_+,\mathbb{R}^\Theta):\\ \mathbf{y}(0)=y,\,\mathbf{y}(t)=y'}} I^*_y(\mathbf{y})\,.$$

Given $\nu(\rho) = (\nu_\theta(\rho)\,, \theta \in \Theta)$, as defined in (1.2), and $y \in \mathbb{S}_{|\Theta|}$, let

$$\Phi(\nu(\rho), y) = \lim_{t\to\infty} \Phi_t(\nu(\rho), y) = \inf_{\substack{\mathbf{y}\in\mathbb{C}(\mathbb{R}_+,\mathbb{R}^\Theta):\\ \mathbf{y}(0)=\nu(\rho),\,\mathbf{y}(t)=y \text{ for some } t}} I^*_{\nu(\rho)}(\mathbf{y})\,.$$

(The limit exists because the infima monotonically decrease with t, as sitting at $\nu(\rho)$ "costs" nothing.) Let A denote the set of solutions ρ of (1.2) and (1.3). For $\rho \in A$, let $G(\rho)$ denote the set of directed graphs that are in-trees with root ρ on the vertex set A. Thus, for every $\rho' \in A$ and $q \in G(\rho)$, there is a unique directed path from ρ' to ρ in q. For $q \in G(\rho)$, let $E(q)$ denote the set of edges of q. Define

$$\begin{aligned} J(\nu(\rho)) = \inf_{q\in G(\rho)} &\sum_{(\rho',\rho'')\in E(q)} \Phi(\nu(\rho'), \nu(\rho'')) \\ &- \inf_{\tilde\rho\in A} \inf_{q\in G(\tilde\rho)} \sum_{(\rho',\rho'')\in E(q)} \Phi(\nu(\rho'), \nu(\rho''))\,. \end{aligned} \tag{3.1}$$

Let the simplex $\mathbb{S}_\Theta$ be endowed with the subspace topology.

Theorem 3.1. *Suppose that the equations in* (1.2) *and* (1.3) *admit finitely many solutions* $\rho = (\rho_1, \dots, \rho_K)$. *Then, the measures* $\mu^{(n)}$ *satisfy an LDP in* $\mathbb{S}_{|\Theta|}$ *for the topology of weak convergence with a continuous deviation function*

$$J(y) = \inf_{\rho\in A} (J(\nu(\rho)) + \Phi(\nu(\rho), y))\,. \tag{3.2}$$

Proof. The proof is done by applying Theorem 2.1 in Puhalskii [9]. Since the set $\mathbb{S}_{|\Theta|}$ is compact so that the measures $\mu^{(n)}$ are exponentially tight and $I^*(\mathbf{y}) = \infty$ unless $\mathbf{y} \in \mathbb{C}(\mathbb{R}_+, \mathbb{S}_{|\Theta|})$, one needs to check the following requirements:

1. if $y^{(n)} \to y$, then the distributions of $Y^{(n)}$ satisfy an LDP with I^*_y,
2. the function $I^*_y(\mathbf{y})$ is lower semicontinuous in $(y, \mathbf{y})$ and the set $\cup_{y\in\mathbb{S}_{|\Theta|}}\{\mathbf{y} \in \mathbb{C}(\mathbb{R}_+, \mathbb{S}_{|\Theta|}) : I^*_y(\mathbf{y}) \le \eta\}$ is compact, for all $\eta \ge 0$,
3. for all $\mathbf{y} \in \mathbb{C}(\mathbb{R}_+, \mathbb{S}_{|\Theta|})$,

$$I^*_y(\mathbf{y}) = \inf_{\mathbf{y}'\in p_s^{-1}(p_s\mathbf{y})} I^*_y(\mathbf{y}') + I^*_{\mathbf{y}(s)}(\vartheta_s\mathbf{y})\,,$$

4. (a) if $I_y^*(\mathbf{y})=0$, then $\inf_{t\geq 0} d(\mathbf{y}(t),A)=0$, where d is a metric on $\mathbb{S}_{|\Theta|}$,
   (b) if $\mathbf{y}(t)=\nu(\rho)$, for all $t\geq 0$, then $I^*_{\nu(\rho)}(\mathbf{y})=0$, where $\rho\in A$,
   (c) for any $\rho,\tilde\rho\in A$, there exists $t>0$ such that $\Phi_t(\nu(\rho),\nu(\tilde\rho))<\infty$,
   (d) for any $\varepsilon>0$, there exists $\eta>0$ such that if $d(y,A)<\eta$, then $\Phi_{s_0}(y,\nu(\rho))<\varepsilon$ and $\Phi_{s_1}(\nu(\rho),y)<\varepsilon$, for some $s_0>0$, $s_1>0$ and $\rho\in A$,
   (e) for any $y\in\mathbb{S}_{|\Theta|}$ and $\varepsilon>0$, there exist $\eta>0$, t_0 and t_1 such that $\Phi_{t_0}(y,\tilde y)<\varepsilon$ and $\Phi_{t_1}(\tilde y,y)<\varepsilon$ provided $d(y,\tilde y)<\eta$.

Part 1 holds by Theorem 2.1. Part 2 is a consequence of Young's product inequality: by (2.2), for $s\leq t$, $\lambda\in\mathbb{R}^\Theta$, and $\varepsilon>0$,

$$\lambda\cdot(\mathbf{y}(t)-\mathbf{y}(s))\leq\varepsilon\int\limits_s^t L(\mathbf{y}(u),\dot{\mathbf{y}}(u))\,du+\varepsilon\int\limits_s^t H(\mathbf{y}(u),\frac{\lambda}{\varepsilon})\,du\,.$$

As $t-s\to 0$, with s and t being bounded, the second term on the righthand side goes to 0 uniformly over $\mathbf{y}$ and over λ from a bounded set. The first term is bounded above by $\varepsilon\int_0^\infty L(\mathbf{y}(u),\dot{\mathbf{y}}(u))\,du\leq\varepsilon\eta$, so, it can be made small uniformly over s and t. The needed property holds by Arzela–Ascoli's theorem. For part 3, note that, by (2.3), $\inf_{\mathbf{y}'\in p_s^{-1}(p_s\mathbf{y})} I_y^*(\mathbf{y}')=\int_0^s L(\mathbf{y}(t),\dot{\mathbf{y}}(t))\,dt$ and $I^*_{\mathbf{y}(s)}(\vartheta_s\mathbf{y})=\int_0^\infty L(\vartheta_s\mathbf{y}(t),d/dt\,(\vartheta_s\mathbf{y}(t)))\,dt=\int_s^\infty L(\mathbf{y}(t),\dot{\mathbf{y}}(t))\,dt$.

As for part 4, Proposition 4 in Antunes et al. [2] implies that if $\mathbf{y}$ satisfies (1.1), then $\mathbf{y}(t)$ converges, as $t\to\infty$, to the set A, which verifies the requirement of part 4(a). Part 4(b), essentially, is about the definition of $\nu(\rho)$. For part 4(c), one can take $t=1$ and $\mathbf{y}(s)=(1-s\wedge 1)\nu(\rho)+s\wedge 1\,\nu(\tilde\rho)$. Part 4(d) is addressed next. Given ρ such that $0<d(y,\nu(\rho))<\eta$, one lets $\mathbf{y}(t)=y+t(\nu(\rho)-y)/d(\nu(\rho),y)$. Then, $\mathbf{y}(0)=y$, $\mathbf{y}(d(\nu(\rho),y))=\nu(\rho)$ and

$$\Phi_{d(\nu(\rho),y)}(y,\nu(\rho))\leq\int\limits_0^{d(\nu(\rho),y)} L(\mathbf{y}(t),\dot{\mathbf{y}}(t))\,dt\,. \tag{3.3}$$

If $t\leq d(\nu(\rho),y)$, then $\mathbf{y}(t)\geq t\nu(\rho)/d(\nu(\rho),y)$ entrywise, so that, on recalling that the set of ρ is finite, there exists $\kappa>0$ such that $\mathbf{y}_\theta(t)\geq t\kappa/d(\nu(\rho),y)$, for all $\theta\in\Theta$. By the definition of $H(y,\lambda)$ in (2.1),

$$H(\mathbf{y}(t),\lambda)\geq\min_k(\alpha_k\wedge\delta_k)\sum_{k=1}^K\Big(\sum_{\theta\in\Theta_k^+}e^{\lambda_{\theta+e_k}-\lambda_\theta}+\sum_{\theta\in\Theta_k^-}e^{\lambda_{\theta-e_k}-\lambda_\theta}\theta_k\Big)\frac{\kappa}{d(\nu(\rho),y)}\,t$$
$$-\sum_{k=1}^K\big(\alpha_k+|\Theta|(\delta_k+\gamma_k)\big)\,. \tag{3.4}$$

Let $\lambda_{\tilde{\theta}} = \min_{\theta\in\Theta}\lambda_\theta$ and $\lambda_{\hat{\theta}} = \max_{\theta\in\Theta}\lambda_\theta$. Let $\tilde{\theta} = \theta_0, \theta_1, \ldots, \theta_\ell = \hat{\theta}$ represent a path from $\tilde{\theta}$ to $\hat{\theta}$, as defined in Remark 2.3. By Jensen's inequality,

$$\sum_{k=1}^{K}\sum_{\theta\in\Theta_k^\pm} e^{\lambda_{\theta\pm e_k}-\lambda_\theta} \geq \sum_{i=1}^{\ell} e^{\lambda_{\theta_i}-\lambda_{\theta_{i-1}}} \geq \ell e^{\sum_{i=1}^{\ell}(\lambda_{\theta_i}-\lambda_{\theta_{i-1}})/\ell} = \ell e^{(\lambda_{\hat{\theta}}-\lambda_{\tilde{\theta}})/\ell} .$$

Assuming that $\lambda_{\tilde{\theta}} < \lambda_{\hat{\theta}}$ so that $\ell \geq 1$ obtains that

$$\begin{aligned}\sum_{\theta\in\Theta}\lambda_\theta\dot{\mathbf{y}}_\theta(t) - H(\mathbf{y}(t),\lambda) \leq |\Theta|(\lambda_{\hat{\theta}} - \lambda_{\tilde{\theta}})&\frac{\max_{\theta\in\Theta}|\nu_\theta(\rho) - y_\theta|}{d(\nu(\rho),y)} \\ &- \min_k(\alpha_k\wedge\delta_k)e^{(\lambda_{\hat{\theta}}-\lambda_{\tilde{\theta}})/|\Theta|}\frac{\kappa}{d(\nu(\rho),y)}\,t \\ &+ \sum_{k=1}^{K}\big(\alpha_k + |\Theta|(\delta_k+\gamma_k)\big) .\end{aligned}$$

A similar inequality holds if all the λ_θ in (3.4) are the same. Maximisation over $\lambda_{\hat{\theta}} - \lambda_{\tilde{\theta}}$ shows that the latter righthand side is bounded above by $d_1 + d_2\ln(d(\nu(\rho),y)/t)$, for suitable constants d_1 and d_2 . Hence, the integral on the right of (3.3) converges to zero as $\eta \to 0$. The argument for $\Phi_{s_1}(\nu(\rho),y)$ is similar: one introduces $\mathbf{y}(t) = \nu(\rho) + t(y-\nu(\rho))/d(\nu(\rho),y)$, notes that $\mathbf{y}_\theta(t) \geq (1-t)\nu_\theta(\rho)/d(\nu(\rho),y)$ and uses a similar bound to (3.4). The checking of part 4(e) is done analogously. □

Remark 3.1. It is noteworthy that if $\Phi(\nu(\rho'),\nu(\rho'')) = 0$, for some ρ' , ρ'' , then ρ' may be omitted in (3.2).

Remark 3.2. Interestingly enough, the quantities $J(\nu(\rho))\,,\rho \in A$, are unique solutions to the system of the balance equations that, for any partition $\{A', A''\}$ of A ,

$$\inf_{\rho'\in A'}\inf_{\rho''\in A''}\big(J(\nu(\rho'))+\Phi(\nu(\rho'),\nu(\rho''))\big) = \inf_{\rho'\in A'}\inf_{\rho''\in A''}\big(J(\nu(\rho''))+\Phi(\nu(\rho''),\nu(\rho'))\big)$$

subject to the normalisation condition that $\inf_{\rho\in A} J(\nu(\rho)) = 0$, see Puhalskii [9].

The next result concerns metastability. It is in the spirit of Freidlin and Wentzell [6], see also Shwartz and Weiss [10]. It is also similar to Corollary 3.1 in Tibi [12], where a proof is outlined assuming a trajectorial LDP. As the argument in Tibi [12] depends on certain contentions in Freidlin and Wentzell [6] being true whose proofs are not available in the literature, a self–contained proof of Theorem 3.2 is provided in the appendix. As before, P_y and E_y denote probability and expectation, respectively, that correspond to the initial condition $Y^{(n)}(0) = y$.

Theorem 3.2. *Let $\nu(\rho)$ be an equilibrium of (1.1) and let D be an open subset of $\mathbb{S}_\Theta$, which contains $\nu(\rho)$. Suppose that the solutions of (1.1) with initial conditions in some neighbourhood of D converge to $\nu(\rho)$ and stay in D when started in D. Let $\tau^{(n)} = \inf\{t \ge 0 : Y^{(n)}(t) \notin D\}$. Let $y^{(n)} \in D \cap S_\Theta^{(n)}$. If $y^{(n)} \to y \in D$, as $n \to \infty$, then*

$$P_{y^{(n)}}\bigl(|\frac{1}{n}\ln\tau^{(n)} - U| > \kappa\bigr) \to 0$$

and

$$\frac{1}{n}\ln E_{y^{(n)}}(\tau^{(n)})^m \to mU\,,$$

where $m \in \mathbb{N}$,

$$U = \inf_{y' \notin D} \Phi(\nu(\rho), y')$$

and $\kappa > 0$ is otherwise arbitrary.

In Antunes et al. [2] stability of equilibria is tackled via the Lyapunov function

$$g(y) = \sum_{\theta\in\Theta} y_\theta \ln(\prod_{k=1}^K \theta_k!\, y_\theta) - \sum_{k=1}^K \frac{\delta_k+\gamma_k}{\gamma_k}\bigl(u\ln u - u\bigr)\Big|_{u=\alpha_k/(\delta_k+\gamma_k)}^{u=(\alpha_k+\gamma_k\sum_{\theta\in\Theta}\theta_k y_\theta)/(\delta_k+\gamma_k)}.$$

In the interior of $\mathbb{S}_\Theta$, see Antunes et al. [2],

$$\nabla g(y)\cdot V(y) = \sum_{k=1}^K \sum_{\theta\in\Theta}((\delta_k+\gamma_k)\theta_k y_\theta - b_k(y))\ln\frac{b_k(y)}{(\delta_k+\gamma_k)\theta_k y_\theta}\,,$$

where

$$b_k(y) = \alpha_k + \gamma_k \sum_{\theta\in\Theta}\theta_k y_\theta\,.$$

Hence, $\nabla g(y)\cdot V(y) \le 0$ so that $g(\mathbf{y}(t))$ is nonincreasing with t along solutions of (1.1) and $\nabla g(y)\cdot V(y) < 0$ provided y is not an equilibrium of (1.1). Furthermore, y is an equilibrium of (1.1) if and only if the differential of g, as a function on $\mathbb{S}_{|\Theta|}$, is zero at y: $dg_{\mathbb{S}_{|\Theta|}}(y) = 0$. If y is a local minimum of g, it is an asymptotically stable equilibrium. In order "to reduce dimension", Antunes et al. [2] introduce the function

$$\phi(\rho) = -\ln Z(\rho) + \sum_{k=1}^K\bigl(\frac{\gamma_k+\delta_k}{\gamma_k}\rho_k - \frac{\alpha_k}{\gamma_k}\ln\rho_k\bigr)\,.$$

By Theorem 3 in Antunes et al. [2], $\rho \in \mathbb{R}_+^K$ is a local minimum of ϕ if and only if $\nu(\rho)$ is a local minimum of g; if ρ is a saddle point of ϕ, then $\nu(\rho)$ is a saddle point of g. Besides, a calculation shows that if $\nu(\rho)$ is an equilibrium, then

$$g(\nu(\rho)) = \phi(\rho) + \sum_{k=1}^K \frac{\alpha_k}{\gamma_k}\bigl(\ln\frac{\alpha_k}{\gamma_k+\delta_k} - 1\bigr)\,.$$

An example of bistability along the lines of the one in Antunes et al. [2] is analysed next. Let $K = 2$. The polynomial equations for (ρ_1, ρ_2) in (1.2) and (1.3) have finitely many solutions by Bézout's theorem as the polynomials in the two equations are coprime, see, e.g., Cox, Little and O'Shea [4]. Antunes et al. [2] show that, for a certain choice of parameters there are at least two stable equilibria. Suppose that class 1 customers require one unit of capacity, so, $A_1 = 1$ whereas class 2 customers require the whole capacity, so, $A_2 = C$. Accordingly, class 1 and class 2 customers cannot coexist at the same node. It stands to reason that there could be two stable states where class 1 customers are prevalent or class 2 customers are prevalent, respectively. This is substantiated next.

By hypotheses, θ_1 takes values in the set $\{0, 1, \ldots, C\}$ and θ_2 takes values in $\{0, 1\}$. Then, with $\rho = (\rho_1, \rho_2)$,

$$Z(\rho) = \sum_{i=0}^{C} \frac{\rho_1^i}{i!} + \rho_2 \tag{3.5}$$

and

$$\phi(\rho) = -\ln\Bigl(\sum_{i=0}^{C} \frac{\rho_1^i}{i!} + \rho_2\Bigr) + \frac{\gamma_1 + \delta_1}{\gamma_1}\,\rho_1 + \frac{\gamma_2 + \delta_2}{\gamma_2}\,\rho_2 - \frac{\alpha_1}{\gamma_1}\,\ln \rho_1 - \frac{\alpha_2}{\gamma_2}\,\ln \rho_2\,.$$

By (1.2) and (1.3) with $k = 1$,

$$Z(\rho) = \frac{\gamma_1 \rho_1 (\rho_2 + \rho_1^C / C!)}{\alpha_1 - \delta_1 \rho_1} \tag{3.6}$$

(as $\sum_{\theta\in\Theta} \theta_1 \nu_\theta(\rho) < \rho_1$, $\alpha_1 > \delta_1 \rho_1$). By (1.2) and (1.3) with $k = 2$, ρ_2 satisfies the quadratic equation

$$\gamma_1(\gamma_2 + \delta_2)\rho_2^2 + \bigl(\gamma_1(\gamma_2 + \delta_2)\frac{\rho_1^C}{C!} - \alpha_2\gamma_1 - \gamma_2(\frac{\alpha_1}{\rho_1} - \delta_1)\bigr)\rho_2 - \alpha_2\gamma_1\,\frac{\rho_1^C}{C!} = 0\,.$$

Hence,

$$\begin{aligned}\rho_2(\rho_1) = {}& \frac{1}{2\gamma_1(\gamma_2 + \delta_2)}\Bigl(-\gamma_1(\gamma_2 + \delta_2)\frac{\rho_1^C}{C!} + \alpha_2\gamma_1 + \gamma_2(\frac{\alpha_1}{\rho_1} - \delta_1) \\ &+ \sqrt{\bigl(\gamma_1(\gamma_2 + \delta_2)\frac{\rho_1^C}{C!} - \alpha_2\gamma_1 - \gamma_2(\frac{\alpha_1}{\rho_1} - \delta_1)\bigr)^2 + 4\gamma_1^2(\gamma_2 + \delta_2)\alpha_2\,\frac{\rho_1^C}{C!}}\;\Bigr)\,.\end{aligned} \tag{3.7}$$

Equating the righthand sides of (3.5) and (3.6) yields

$$h(\rho) = 0\,,$$

where

$$h(\rho) = \sum_{i=0}^{C} \frac{\rho_1^i}{i!} + \rho_2 - \frac{\gamma_1 \rho_1}{\alpha_1 - \delta_1 \rho_1} \big(\frac{\rho_1^C}{C!} + \rho_2 \big) . \tag{3.8}$$

Let $C = 20$, $\alpha_1 = .5$, $\alpha_2 = 9$, $\gamma_1 = \gamma_2 = 1$ and $\delta_1 = \delta_2 = .01$. (The parameters for which Antunes et al. [2] show the existence of two stable equilibria are $C = 20$, $\alpha_1 = .68$, $\alpha_2 = 9$, $\gamma_1 = \gamma_2 = 1$ and $\delta_1 = \delta_2 = 0$. It is of interest to allow nonzero δ_1 and (or) δ_2.) For ρ_1 close to zero, the term α_1/ρ_1 on the right of (3.7) dominates, ρ_2 decreases rapidly as ρ_1 increases, so does the righthand side of (3.8), its first zero being $\rho_1 \approx .5966$. The righthand side of (3.8) keeps decreasing as the leftmost sum starts taking over until it reaches a minimum of approximately -23.556 at $\rho_1 \approx 2.8861$ and begins to increase and crosses the zero level for a second time for $\rho_1 \approx 4.1786$. It keeps growing reaching a maximum of approximately $1.5794 \cdot 10^5$ at $\rho_1 \approx 12.896$ until another change of a dominating term when $-\gamma_1\rho_1/(\alpha_1 - \rho_1\delta_1)\,\rho_1^C/C!$ takes over. The righthand side of (3.8) then plunges to $-\infty$, crossing the zero level for a third time at $\rho_1 \approx 13.72715$ in the process. Graphs in Fig. 1 and Fig. 2 provide an illustration. Thus, all in all, there are three equilibria: $\rho^{(1)} \approx (.5966, 8.8293)$,

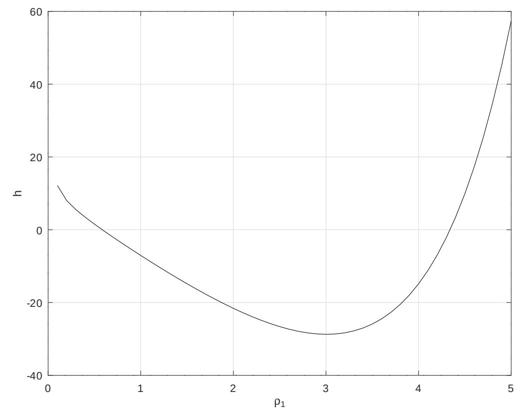

Figure 1. Function $h(\rho_1, \rho_2(\rho_1))$ for ρ_1 small

$\rho^{(2)} \approx (4.1786, 8.1115)$, and $\rho^{(3)} \approx (13.72715, 8.9906)$. The second derivatives of ϕ are

$$\nabla^2 \phi(\rho^{(1)}) \approx \begin{pmatrix} 1.2633 & 0.016025 \\ 0.016025 & 0.1243 \end{pmatrix},$$

$$\nabla^2 \phi(\rho^{(3)}) \approx \begin{pmatrix} 0.015412 & 1.1085 \cdot 10^{-6} \\ 1.1085 \cdot 10^{-6} & 0.1113 \end{pmatrix},$$

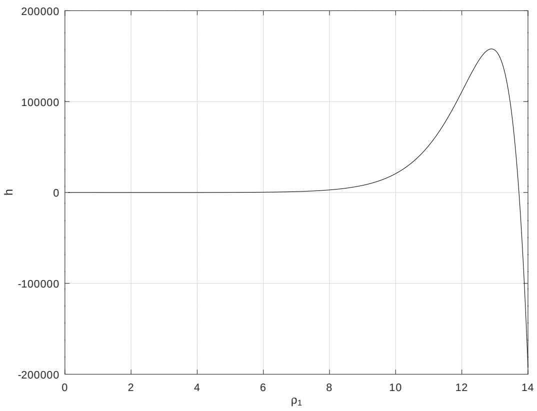

Figure 2. Function $h(\rho_1, \rho_2(\rho_1))$ globally

and

$$\nabla^2\phi(\rho^{(2)}) \approx \begin{pmatrix} -0.069679 & 0.0121200 \\ 0.012120 & 0.1370 \end{pmatrix},$$

the eigenvalues in the latter case being -0.070387 and 0.137708 approximately. With $\nabla^2\phi(\rho^{(1)})$ and $\nabla^2\phi(\rho^{(3)})$ being positive definite, $\rho^{(1)}$ and $\rho^{(3)}$ are local minima, whereas $\rho^{(2)}$ is a saddle point. A 3D mesh plot of $\phi(\rho)$ with a contour plot underneath is in Fig. 3. Therefore, $\nu(\rho^{(1)}))$ and $\nu(\rho^{(3)})$ are asymptotically

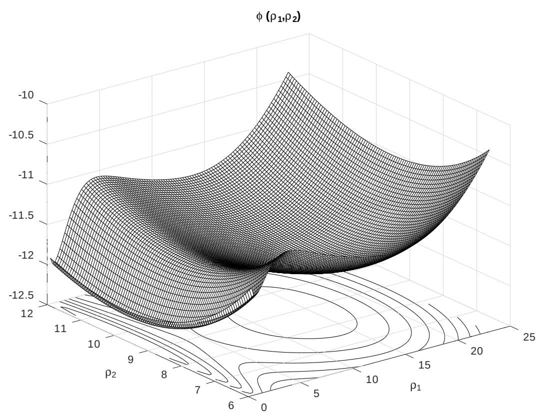

Figure 3. 3D mesh plot and contour plot of $\phi(\rho_1, \rho_2)$

stable equilibria so that the network process spends exponentially long periods of time in the neighbourhoods of those equilibria, while $\nu(\rho^{(2)})$ is an unstable equilibrium.

The expected number of class k customers being

$$EQ_k = \rho_k(1 - \sum_{\theta:\, \sum_{k'} A_{k'}\theta_{k'} > C - A_k} \nu_\theta(\rho))$$

implies that the average numbers of class 1 and class 2 customers are

$$EQ_1 = \rho_1(1 - \frac{1}{Z(\rho)}\frac{\rho_1^C}{C!})$$

and

$$EQ_2 = \rho_2(1 - \frac{1}{Z(\rho)}(\sum_{\theta_1=1}^{C}\frac{\rho_1^{\theta_1}}{\theta_1!} + \rho_2)) = \frac{\rho_2}{Z(\rho)}\,,$$

respectively. For the stable equilibria, calculations yield $(EQ_1^{(1)}, EQ_2^{(1)}) \approx (.5966, .8281)$ and $(EQ_1^{(3)}, EQ_2^{(3)}) \approx (13.365, 1.0235 \cdot 10^{-5})$. Thus, for $\rho = \rho^{(1)}$, class 2 customers are prevalent and, for $\rho = \rho^{(3)}$, class 1 customers are prevalent. (The pattern of $h(\rho)$ first decreasing, then increasing and decreasing again is sensitive to the values of δ_1 and δ_2. When $\delta_1 = \delta_2 = .1$, only a downward trend is present, so, there is only one equilibrium.)

Also, calculations yield $\phi(\rho^{(1)}) = -12.284$, $\phi(\rho^{(2)}) = -11.560$, $\phi(\rho^{(3)}) = -12.043$ so that $g(\nu(\rho^{(2)})) > g(\nu(\rho^{(1)})) \vee g(\nu(\rho^{(3)}))$.Thus, if $\mathbf{y}(0) = \nu(\rho^{(2)})$ experiences a small displacement Δ at time zero in a direction collinear with the direction of the eigenvector with a negative eigenvalue, then $g(y+\Delta) < g(y)$, so that the associated trajectory $\mathbf{y}(t)$ will end up in one of the equilibria $\nu(\rho^{(1)})$ or $\nu(\rho^{(3)})$. Hence, either $\Phi(\nu(\rho^{(2)}), \nu(\rho^{(1)}) = 0$ or $\Phi(\nu(\rho^{(2)}), \nu(\rho^{(3)})) = 0$.

On denoting $\nu^{(i)} = \nu(\rho^{(i)})$, by (3.1),

$$\begin{aligned}
J(\nu^{(1)}) &= (\Phi(\nu^{(2)}, \nu^{(1)}) + \Phi(\nu^{(3)}, \nu^{(2)})) \wedge (\Phi(\nu^{(2)}, \nu^{(1)}) + \Phi(\nu^{(3)}, \nu^{(1)})) \\
&\qquad \wedge(\Phi(\nu^{(3)}, \nu^{(1)}) + \Phi(\nu^{(2)}, \nu^{(3)})) - \Psi\,, \\
J(\nu^{(2)}) &= (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(3)}, \nu^{(1)})) \wedge (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(3)}, \nu^{(2)})) \\
&\qquad \wedge(\Phi(\nu^{(3)}, \nu^{(2)}) + \Phi(\nu^{(1)}, \nu^{(3)})) - \Psi\,, \\
J(\nu^{(3)}) &= (\Phi(\nu^{(2)}, \nu^{(3)}) + \Phi(\nu^{(1)}, \nu^{(2)})) \wedge (\Phi(\nu^{(2)}, \nu^{(3)}) + \Phi(\nu^{(1)}, \nu^{(3)})) \\
&\qquad \wedge(\Phi(\nu^{(1)}, \nu^{(3)}) + \Phi(\nu^{(2)}, \nu^{(1)})) - \Psi\,,
\end{aligned}$$

where

$$\begin{aligned}
\Psi = {} & (\Phi(\nu^{(2)}, \nu^{(1)}) + \Phi(\nu^{(3)}, \nu^{(2)})) \wedge (\Phi(\nu^{(2)}, \nu^{(1)}) + \Phi(\nu^{(3)}, \nu^{(1)})) \\
& \wedge (\Phi(\nu^{(3)}, \nu^{(1)}) + \Phi(\nu^{(2)}, \nu^{(3)})) \wedge (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(3)}, \nu^{(1)}))
\end{aligned}$$

$$\wedge\,(\Phi(\nu^{(1)},\nu^{(2)})+\Phi(\nu^{(3)},\nu^{(2)}))\wedge(\Phi(\nu^{(3)},\nu^{(2)})+\Phi(\nu^{(1)},\nu^{(3)}))$$
$$\wedge\,(\Phi(\nu^{(2)},\nu^{(3)})+\Phi(\nu^{(1)},\nu^{(2)}))\wedge(\Phi(\nu^{(2)},\nu^{(3)})+\Phi(\nu^{(1)},\nu^{(3)}))$$
$$\wedge\,(\Phi(\nu^{(1)},\nu^{(3)})+\Phi(\nu^{(2)},\nu^{(1)}))\,.$$

As a consequence, if $\Phi(\nu^{(2)},\nu^{(1)})=0$, then

$$J(\nu^{(1)})=\Phi(\nu^{(3)},\nu^{(2)})\wedge\Phi(\nu^{(3)},\nu^{(1)})-\Psi\,,$$
$$J(\nu^{(2)})=(\Phi(\nu^{(1)},\nu^{(2)})+\Phi(\nu^{(3)},\nu^{(1)}))\wedge(\Phi(\nu^{(1)},\nu^{(2)})+\Phi(\nu^{(3)},\nu^{(2)}))$$
$$\wedge(\Phi(\nu^{(3)},\nu^{(2)})+\Phi(\nu^{(1)},\nu^{(3)}))-\Psi\,,$$
$$J(\nu^{(3)})=(\Phi(\nu^{(2)},\nu^{(3)})+\Phi(\nu^{(1)},\nu^{(2)}))\wedge\Phi(\nu^{(1)},\nu^{(3)})-\Psi$$

and

$$\Psi=\Phi(\nu^{(3)},\nu^{(2)})\wedge\Phi(\nu^{(3)},\nu^{(1)})\wedge(\Phi(\nu^{(2)},\nu^{(3)})+\Phi(\nu^{(1)},\nu^{(2)}))\wedge\Phi(\nu^{(1)},\nu^{(3)})\,.$$

It is being conjectured that, furthermore, $\Phi(\nu^{(2)},\nu^{(1)})=\Phi(\nu^{(2)},\nu^{(3)})=0$ and that $\Phi(\nu^{(3)},\nu^{(2)})\le\Phi(\nu^{(3)},\nu^{(1)})$ and $\Phi(\nu^{(1)},\nu^{(2)})\le\Phi(\nu^{(1)},\nu^{(3)})$, in which case the expressions above simplify.

A. Proof of Theorem 3.2

The proof is along the lines of the developments in Freidlin and Wentzell [6] and Shwartz and Weiss [10]. Suppose, it has been proved that

$$\limsup_{n\to\infty}\frac{1}{n}\,\ln E_{y^{(n)}}(\tau^{(n)})^m\le mU \tag{A.1}$$

and

$$\lim_{n\to\infty}P_{y^{(n)}}\big(\frac{1}{n}\,\ln\tau^{(n)}\le U-3\kappa\big)=0\,. \tag{A.2}$$

By Markov's inequality and (A.1) with $m=1$,

$$\limsup_{n\to\infty}P_{y^{(n)}}\big(\frac{1}{n}\,\ln\tau^{(n)}\ge U+\kappa\big)^{1/n}\le\limsup_{n\to\infty}(E\tau^{(n)})^{1/n}e^{-(U+\kappa)}\le e^{-\kappa}$$

so that

$$\lim_{n\to\infty}P_{y^{(n)}}\big(\frac{1}{n}\,\ln\tau^{(n)}\ge U+\kappa\big)=0\,.$$

By Jensen's inequality, for $\varepsilon>0$,

$$\frac{1}{n}\,\ln E_{y^{(n)}}(\tau^{(n)}\vee\varepsilon)^m\ge\frac{m}{n}\,E_{y^{(n)}}(\ln\tau^{(n)}\vee\ln\varepsilon)\,.$$

By (A.2) and Fatou's lemma,

$$\liminf_{n\to\infty} E_{y^{(n)}} \frac{1}{n} (\ln \tau^{(n)} \vee \ln \varepsilon) \geq U\,.$$

On the other hand,

$$\begin{aligned}\ln E_{y^{(n)}} (\tau^{(n)} \vee \varepsilon)^m &\leq \ln(E_{y^{(n)}} (\tau^{(n)})^m + \varepsilon^m) \\ &\leq \ln 2 + \ln(E_{y^{(n)}} (\tau^{(n)})^m) \vee \ln \varepsilon^m\,.\end{aligned}$$

Hence,

$$\liminf_{n\to\infty} \frac{1}{n} \ln E_{y^{(n)}} (\tau^{(n)})^m \geq mU\,.$$

Next, (A.1) and (A.2) are proved.

For $r > 0$, let B_r denote the open ball of radius r about $\nu(\rho)$. Let $\eta > 0$ be small enough for $B_{3\eta}$ to belong to D. Let $T(y')$ denote the length of time that it takes the solution $\mathbf{y}$ of (1.1) with y' as an initial point to reach $\mathrm{cl}B_{\eta/2}$, where $y' \in \mathrm{cl}D$. The function $T(y')$ is upper semicontinuous. Since D is bounded, so is $T(y')$. Let $T_1 = \max_{y'\in \mathrm{cl}\, D} T(y')$. Let $\sigma^{(n)}$ represent the first time when $Y^{(n)}$ reaches the closed ball $\mathrm{cl}B_\eta$. Let $\mathbf{y}(t)$ solve (1.1) with $\mathbf{y}(0) = y$. Since $y^{(n)} \to y$ and $Y^n(t) \to \mathbf{y}(t)$ uniformly on $[0, T_1]$ in $P_{y^{(n)}}$–probability, it may be assumed that

$$P_{y^{(n)}}(\sigma^{(n)} \leq T_1) \geq \frac{1}{2}\,, \tag{A.3}$$

provided n is great enough.

Lemma 3.1 in Puhalskii [9] implies that $\Phi(\nu(\rho), y')$ is a continuous function of y'. Therefore,

$$U = \inf_{y'\notin \mathrm{cl}\, D} \Phi(\nu(\rho), y')\,. \tag{A.4}$$

Let D_η represent the open η–neighbourhood of D and let

$$U_\eta = \inf_{y\notin D_\eta} \Phi(\nu(\rho), y)\,.$$

By (A.4), given $\beta > 0$, one may assume that η is small enough so that $U_\eta \leq U + \beta$. (One can assume that $U < \infty$.) Furthermore, given $y'^{(n)} \in B_\eta$, there exist $\mathbf{y}^{(n)}$ and $t^{(n)}$ such that $\mathbf{y}^{(n)}(0) = y'^{(n)}$, $\mathbf{y}^{(n)}(t^{(n)}) \notin D_\eta$ and $I^*_{y'^{(n)}}(\mathbf{y}^{(n)}) \leq U_\eta + \beta$. (Note that one can get from $y'^{(n)}$ to $\nu(\rho)$ at an arbitrarily small cost by following the solution of (1.1).)

It is shown next that the sequence $t^{(n)}$ can be chosen bounded. Firstly, $t^{(n)}$ can be chosen as the smallest t with $\mathbf{y}^{(n)}(t) \notin D_\eta$. Let $s^{(n)}$ be the last time t such that $\mathbf{y}^{(n)}(t) \in \mathrm{cl}B_\eta$. Then, for $t \in [s^{(n)}, t^{(n)}]$, the function $\mathbf{y}^{(n)}(t)$ takes values in the closed set $\mathrm{cl}D_\eta \setminus B_\eta$. Let T denote the maximal length of time it takes a solution of (1.1) with an initial point in $\mathrm{cl}D_\eta \setminus B_\eta$ to get to $\mathrm{cl}B_\eta$. If

$t^{(n)} - s^{(n)} \leq NT$, for some N, the proof is over. Otherwise, on denoting, for $W \subset \mathbb{C}(\mathbb{R}_+, \mathbb{R}^{|\Theta|})$, $I^*(W) = \inf_{\mathbf{y}\in W} I^*(\mathbf{y})$, by (2.3),

$$\begin{aligned}
I^*_{y'^{(n)}}(\mathbf{y}^{(n)}) &\geq \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(\vartheta_{s^{(n)}}\mathbf{y}^{(n)})\\
&= \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} \big(I^*_{y'}(p_T^{-1}(p_T(\vartheta_{s^{(n)}}\mathbf{y}^{(n)}))) + I^*_{\mathbf{y}^{(n)}(s^{(n)}+T)}(\vartheta_{s^{(n)}+T}\mathbf{y}^{(n)})\big)\\
&\geq \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(p_T^{-1}(p_T(\vartheta_{s^{(n)}}\mathbf{y}^{(n)})))\\
&\quad + \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(\vartheta_{s^{(n)}+T}\mathbf{y}^{(n)})\\
&= \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(p_T^{-1}(p_T(\vartheta_{s^{(n)}}\mathbf{y}^{(n)})))\\
&\quad + \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} (I^*_{y'}(p_T^{-1}(p_T\vartheta_{s^{(n)}+T}\mathbf{y}^{(n)}))\\
&\quad + I^*_{\mathbf{y}^{(n)}(s^{(n)}+2T)}(\vartheta_{s^{(n)}+2T}\mathbf{y}^{(n)}))\\
&\geq \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(p_T^{-1}(p_T(\vartheta_{s^{(n)}}\mathbf{y}^{(n)})))\\
&\quad + \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(p_T^{-1}(p_T\vartheta_{s^{(n)}+T}\mathbf{y}^{(n)}))\\
&\quad + \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(\vartheta_{s^{(n)}+2T}\mathbf{y}^{(n)})\,.
\end{aligned}$$

Continuing on yields, for arbitrary $N \in \mathbb{N}$ such that $t^{(n)} - s^{(n)} \geq NT$,

$$\begin{aligned}
I^*_{y'^{(n)}}(\mathbf{y}^{(n)}) &\geq \sum_{m=0}^{N} \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} I^*_{y'}(p_T^{-1}(p_T\vartheta_{s^{(n)}+mT}\mathbf{y}^{(n)}))\\
&\geq N \inf_{y'\in \mathrm{cl}D_\eta\setminus B_\eta} \inf_{\substack{\mathbf{y}' :\, \mathbf{y}'(t)\in \mathrm{cl}D_\eta\setminus B_\eta\\ \text{for all } t\in[0,T]}} I^*_{y'}(\mathbf{y}')\,.
\end{aligned}$$

The latter infimum is positive because no solution of (1.1) belongs to the set $\{\mathbf{y}' : \mathbf{y}'(t) \in \mathrm{cl}D_\eta \setminus B_\eta \text{ for all } t \in [0,T]\}$. It follows that the values of N have to be bounded, so, the $t^{(n)} - s^{(n)}$ have to be bounded.

The sequence $s^{(n)}$ can be assumed bounded. Indeed, one can change $\mathbf{y}^{(n)}$ by replacing the piece of $\mathbf{y}^{(n)}$ on $[0, s^{(n)}]$ with a straight line segment connecting $y'^{(n)}$ and $\mathbf{y}^{(n)}(s^{(n)})$. The modified trajectory is given by $\tilde{\mathbf{y}}^{(n)}(t) = y'^{(n)} + t(\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)})/|\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)}|$, for $t \in [0, |\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)}|]$ and $\tilde{\mathbf{y}}^{(n)}(t) = \mathbf{y}^{(n)}(t - |\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)}| + s^{(n)})$, for $t \geq |\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)}|$. The last time $\tilde{\mathbf{y}}^{(n)}$ visits $\mathrm{cl}B_\eta$ is $\tilde{s}^{(n)} = |\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)}|$. The $\tilde{s}^{(n)}$ are thus bounded. With $\tilde{t}^{(n)} = t^{(n)} - s^{(n)} + |\mathbf{y}^{(n)}(s^{(n)}) - y'^{(n)}|$, $\tilde{t}^{(n)}$ is the smallest t with $\tilde{\mathbf{y}}^{(n)}(t) \notin D_\eta$. In addition, it is possible to choose η small enough to ensure that $I^*_{y'^{(n)}}(\tilde{\mathbf{y}}^{(n)}) \leq I^*_{y'^{(n)}}(\mathbf{y}^{(n)}) + \beta$.

Since $y'^{(n)}(t^{(n)}) \notin D_\eta$, $Y^{(n)}(t^{(n)}) \notin D$ provided $|Y^{(n)}(t^{(n)}) - \mathbf{y}^{(n)}(t^{(n)})| < \eta$ so that

$$P_{y'^{(n)}}(\tau^{(n)} \leq t^{(n)}) \geq P_{y'^{(n)}}(\sup_{s \leq t^{(n)}} |Y^{(n)}(s) - \mathbf{y}^{(n)}(s)| < \eta)\,.$$

Denote $D^{(n)} = D \cap \mathbb{S}_\Theta^{(n)}$ and $B_r^{(n)} = B_r \cap \mathbb{S}_\Theta^{(n)}$. Assuming that $t^{(n)} \leq T_2$, that $y'^{(n)} \in B_\eta^{(n)}$, that $y'^{(n)} \to y'$, that $t^{(n)} \to \hat{t}$ and that $\mathbf{y}^{(n)} \to \hat{\mathbf{y}}$ yield

$$\begin{aligned}
\liminf_{n\to\infty} \frac{1}{n} \ln P_{y'^{(n)}}(\tau^{(n)} \leq T_2) &\geq \liminf_{n\to\infty} \frac{1}{n} \ln P_{y'^{(n)}}(\tau^{(n)} \leq t^{(n)}) \\
&\geq \liminf_{n\to\infty} \frac{1}{n} \ln P_{y'^{(n)}}(\sup_{s \leq t^{(n)}} |Y^{(n)}(s) - \mathbf{y}^{(n)}(s)| < \eta) \\
&\geq -\inf\{I^*_{y',\hat{t}}(\mathbf{y}) : \sup_{s \leq \hat{t}} |\mathbf{y}(s) - \hat{\mathbf{y}}(s)| < \eta\} \\
&\geq -I^*_{y',\hat{t}}(\hat{\mathbf{y}}) \geq -I^*_{y'}(\hat{\mathbf{y}}) \geq -U_\eta - 2\beta \geq -U - 3\beta\,.
\end{aligned}$$

Since, for arbitrary $y' \in D^{(n)}$,

$$P_{y'}(\tau^{(n)} \leq T_1 + T_2) \geq P_{y'}(\sigma^{(n)} \leq T_1) \inf_{y'' \in B_\eta^{(n)}} P_{y''}(\tau^{(n)} \leq T_2),$$

the argument of the proof of (A.3) yields

$$\liminf_{n\to\infty} \frac{1}{n} \ln \inf_{y' \in D^{(n)}} P_{y'}(\tau^{(n)} \leq T_3) \geq -U\,,$$

where $T_3 = T_1 + T_2$. Thus, for all $\beta > 0$,

$$\sup_{y' \in D^{(n)}} P_{y'}(\tau^{(n)} > T_3) \leq 1 - e^{n(-U-\beta)}\,, \tag{A.5}$$

provided n is great enough. By the Markov property, for $\ell \in \mathbb{N}$,

$$\begin{aligned}
&P_{y'}(\tau^{(n)} > \ell T_3 | \tau^{(n)} > (\ell-1)T_3) \\
&\quad = P_{y'}(Y^{(n)}(t) \in D\,, t \in [(\ell-1)T_3, \ell T_3] | Y^{(n)}(t) \in D\,, t \in [0, (\ell-1)T_3]) \\
&\quad = P_{y'}(Y^{(n)}(t) \in D\,, t \in [0, T_3] | Y^{(n)}(0) \in D) \\
&\quad \leq \sup_{y'' \in D^{(n)}} P_{y''}(\tau^{(n)} > T_3)\,,
\end{aligned} \tag{A.6}$$

which implies that

$$\sup_{y' \in D^{(n)}} P_{y'}(\tau^{(n)} > \ell T_3) \leq \sup_{y' \in D^{(n)}} P_{y'}(\tau^{(n)} > T_3)^\ell\,.$$

By (A.5),

$$\sup_{y'\in D^{(n)}} P_{y'}(\tau^{(n)} > \ell T_3) \le (1-e^{n(-U-\beta)})^{\ell}\,.$$

Therefore,

$$\begin{aligned}
E_{y^{(n)}}(\tau^{(n)})^m &= m\int\limits_0^\infty u^{m-1}P_{y^{(n)}}(\tau^{(n)} > u)\,du\\
&\le mT_3\sum_{\ell=0}^\infty((\ell+1)T_3)^{m-1}P_{y^{(n)}}(\tau^{(n)} > \ell T_3)\\
&\le mT_3^m\sum_{\ell=0}^\infty(\ell+1)^{m-1}(1-e^{n(-U-\beta)})^{\ell}\\
&\le mT_3^m(1-e^{n(-U-\beta)})^{-2}\int\limits_0^\infty u^{m-1}(1-e^{n(-U-\beta)})^u\,du\\
&= \frac{m!T_3^m(1-e^{n(-U-\beta)})^{-2}}{(-\ln(1-e^{n(-U-\beta)}))^m}\\
&\le m!T_3^m(1-e^{n(-U-\beta)})^{-2}e^{mn(U+\beta)}
\end{aligned}$$

proving (A.1).

Let $\sigma_0^{(n)} = 0$ and, for $i\in\mathbb{Z}_+$,

$$\tau_i^{(n)} = \inf\{t > \sigma_i^{(n)}\,:\, Y^{(n)}(t)\in B_\eta\}$$

and

$$\sigma_{i+1}^{(n)} = \inf\{t > \tau_i^{(n)}\,:\, Y^{(n)}(t)\notin B_{2\eta}\}\,.$$

One has that

$$\begin{aligned}
&P_{y^{(n)}}(\tau^{(n)} < e^{n(U-3\kappa)})\\
&\quad= \sum_{i=0}^\infty P_{y^{(n)}}(\sigma_i^{(n)} \le \tau^{(n)} < \tau_i^{(n)}, \tau^{(n)} < e^{n(U-3\kappa)})\\
&\quad\le \sum_{i=0}^{\lfloor e^{n(U-2\kappa)}\rfloor} P_{y^{(n)}}(\sigma_i^{(n)} \le \tau^{(n)} < \tau_i^{(n)}) + P_{y^{(n)}}(\tau^{(n)}_{\lfloor e^{n(U-2\kappa)}\rfloor} < e^{n(U-3\kappa)})\\
&\quad= P_{y^{(n)}}(\tau^{(n)} < \tau_0^{(n)})\\
&\qquad+ \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} E_{y^{(n)}}(\mathbf{1}_{\{\sigma_i^{(n)}\le\tau^{(n)}\}}P_{y^{(n)}}(\tau^{(n)} < \tau_i^{(n)}|\mathcal{F}^{(n)}(\sigma_i^{(n)})))
\end{aligned}$$

$$+ P_{y^{(n)}}\Big(\frac{\tau^{(n)}_{\lfloor e^{n(U-2\kappa)}\rfloor}}{e^{n(U-2\kappa)}} < e^{-n\kappa}\Big)\,. \tag{A.7}$$

Since $\tau^{(n)} = \inf\{t \ge \sigma_i^{(n)} :\, Y^{(n)}(t) \notin D\}$ and $\tau_i^{(n)} = \inf\{t \ge \sigma_i^{(n)} :\, Y^{(n)}(t) \in B_\eta\}$ on the event $\{\sigma_i^{(n)} \le \tau^{(n)}\}$, by the strong Markov property, for $i \in \mathbb{N}$, $\ell \in \mathbb{N}$, and n great enough,

$$\begin{aligned}
&\mathbf{1}_{\{\sigma_i^{(n)} \le \tau^{(n)}\}} P_{y^{(n)}}(\tau^{(n)} < \tau_i^{(n)} | \mathcal{F}^{(n)}(\sigma_i^{(n)}))\\
&\quad = \mathbf{1}_{\{\sigma_i^{(n)} \le \tau^{(n)}\}} P_{Y^{(n)}(\sigma_i^{(n)})}(\tau^{(n)} < \tau_0^{(n)})\\
&\quad \le \sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau^{(n)} < \tau_0^{(n)}) \qquad\qquad (A.8)\\
&\quad \le \sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau_0^{(n)} > \ell T_1) + \sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau^{(n)} \le \ell T_1)\,.
\end{aligned}$$

For arbitrary $\tilde y^{(n)} \in B_{3\eta}^{(n)}$ converging to some $\tilde y$, $P_{\tilde y^{(n)}}(\tau_0^{(n)} \le T_1) \to 1$, as $n \to \infty$. Moreover, with $\tilde{\mathbf{y}}$ solving (1.1) for $\tilde{\mathbf{y}}(0) = \tilde y$ and with $\tilde t_1$ being the length of time it takes $\tilde{\mathbf{y}}$ to reach $\operatorname{cl} B_{\eta/2}$,

$$\begin{aligned}
P_{\tilde y^{(n)}}(\tau_0^{(n)} > T_1) &\le P_{\tilde y^{(n)}}(\tau_0^{(n)} > \tilde t_1)\\
&\le P_{\tilde y^{(n)}}(Y^{(n)}(\tilde t_1) \notin B_\eta) \le P_{\tilde y^{(n)}}(|Y^{(n)}(\tilde t_1) - \tilde{\mathbf{y}}(\tilde t_1)| > \frac{\eta}{2})
\end{aligned}$$

so that

$$\limsup_{n\to\infty} \frac{1}{n} \ln P_{\tilde y^{(n)}}(\tau_0^{(n)} > T_1) \le -\inf(I^*_{\tilde y}(\mathbf{y}') :\, |\mathbf{y}'(\tilde t_1) - \tilde{\mathbf{y}}(\tilde t_1)| \ge \frac{\eta}{2}) < 0\,. \tag{A.9}$$

It follows that, for some $\chi > 0$ and all n great enough,

$$\sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau_0^{(n)} > T_1) \le e^{-n\chi}\,.$$

For $\ell \in \mathbb{N}$, in analogy with (A.6),

$$\sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau_0^{(n)} > \ell T_1 | \tau_1^{(n)} > (\ell-1)T_1) \le \sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau_0^{(n)} > T_1)$$

so that

$$\sup_{y' \in B_{3\eta}^{(n)}} P_{y'}(\tau_0^{(n)} > \ell T_1) \le e^{-n\ell\chi}\,. \tag{A.10}$$

The second term on the rightmost side of (A.8) is dealt with next. Let $\breve y^{(n)} \in B_{3\eta}^{(n)}$ converge to $\breve y \in \operatorname{cl} B_{3\eta}$. Then, for some $\breve{\mathbf{y}} \in \{\mathbf{y} :\, \mathbf{y}(0) = \breve y, \mathbf{y}(t) \notin D$ for some $t \in [0, \ell T_1]\}$, the latter set being closed and denoted by F,

$$\limsup_{n\to\infty} \frac{1}{n} \ln P_{\breve y^{(n)}}(\tau^{(n)} \le \ell T_1)$$

$$= \limsup_{n\to\infty} \frac{1}{n} \ln P_{\breve{y}^{(n)}}(Y^{(n)}(t) \notin D \text{ for some } t \le \ell T_1)$$
$$\le -\inf_{\mathbf{y}\in F} I^*_{\breve{y}}(\mathbf{y}) = -I^*_{\breve{y}}(\breve{\mathbf{y}})\,. \tag{A.11}$$

Now, if $\check{\mathbf{y}}$ with $\check{\mathbf{y}}(0) = \nu(\rho)$ is obtained from $\breve{\mathbf{y}}$ by inserting a straight line segment joining points $\nu(\rho)$ and $\breve{\mathbf{y}}(0)$, then, provided η is small enough, $I^*_{\nu(\rho)}(\check{\mathbf{y}}) \le I^*_{\breve{y}}(\breve{\mathbf{y}}) + \kappa/2$ so that, for all n great enough,

$$\limsup_{n\to\infty} \frac{1}{n} \ln P_{\breve{y}^{(n)}}(\tau^{(n)} \le \ell T_1) \le -\inf_{y'\notin D} \Phi(\nu(\rho), y') + \frac{\kappa}{2}\,. \tag{A.12}$$

Thus, for all n great enough,

$$\sup_{y'\in B^{(n)}_{3\eta}} P_{y'}(\tau^{(n)} \le \ell T_1) \le e^{-n(U-\kappa)}\,. \tag{A.13}$$

By (A.8), (A.10) and (A.13), for $i \in \mathbb{N}$, choosing ℓ judiciously, for n great,

$$\mathbf{1}_{\{\sigma_i^{(n)} \le \tau^{(n)}\}} P_{y^{(n)}}(\tau^{(n)} < \tau_i^{(n)} | \mathcal{F}^{(n)}(\sigma_i^{(n)})) \le 2e^{-n(U-\kappa)}\,. \tag{A.14}$$

The first term on the rightmost side of (A.7) is tackled similarly. In analogy with (A.9),

$$\limsup_{n\to\infty} \frac{1}{n} \ln P_{y^{(n)}}(\tau_0^{(n)} > T_1) \le -\inf(I^*_y(\mathbf{y}') : |\mathbf{y}'(\bar{t}_1) - \bar{\mathbf{y}}(\bar{t}_1)| \ge \frac{\eta}{2}) < 0\,,$$

where $\bar{\mathbf{y}}$ solves (1.1) with $\bar{\mathbf{y}}(0) = y$ and $\bar{t}_1$ is the length of time it takes $\bar{\mathbf{y}}$ to hit $\operatorname{cl} B_{\eta/2}$. In analogy with (A.11) and (A.12), with $F' = \{\mathbf{y} : \mathbf{y}(t) \notin D \text{ for some } t \le T_1\}$,

$$\limsup_{n\to\infty} \frac{1}{n} \ln P_{y^{(n)}}(\tau^{(n)} \le T_1) = \limsup_{n\to\infty} \frac{1}{n} \ln P_{y^{(n)}}(Y^{(n)}(t) \notin D \text{ for some } t \le T_1)$$
$$\le -\inf_{\mathbf{y}'\in F'} I^*_y(\mathbf{y}')\,.$$

The set F' being closed, the latter infimum is attained. On the other hand, the solution of (1.1) started at y does not belong to F' as it does not leave D, so, the infimum is less than zero. Hence, $\chi > 0$ can be chosen to satisfy the inequality, for all n great enough, $P_{y^{(n)}}(\tau_0^{(n)} > T_1) \vee P_{y^{(n)}}(\tau^{(n)} \le T_1) \le e^{-n\chi}$. Therefore,

$$P_{y^{(n)}}(\tau^{(n)} < \tau_0^{(n)}) \le P_{y^{(n)}}(\tau_0^{(n)} > T_1) + P_{y^{(n)}}(\tau^{(n)} \le T_1) \le 2e^{-n\chi}\,. \tag{A.15}$$

By (A.7), (A.14), and (A.15),

$$P_{y^{(n)}}(\tau^{(n)} < e^{n(U-3\kappa)}) \le 4e^{-n\kappa} + P_{y^{(n)}}\big(e^{-n(U-2\kappa)} \tau^{(n)}_{\lfloor e^{n(U-2\kappa)}\rfloor}) < e^{-n\kappa}\big)\,. \tag{A.16}$$

Since

$$\tau^{(n)}_{\lfloor e^{n(U-2\kappa)}\rfloor} \geq \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} (\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1\,,$$

by the strong Markov property,

$$\begin{aligned} E_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1 | \mathcal{F}^{(n)}(\sigma_i^{(n)})) &= E_{Y^{(n)}(\sigma_i^{(n)})}(\tau_0^{(n)} \wedge 1) \\ &\geq \inf_{y' \in B_{3\eta}^{(n)} \setminus B_{2\eta}^{(n)}} E_{y'}(\tau_0^{(n)} \wedge 1)\,, \end{aligned}$$

so, assuming n is great enough,

$$\begin{aligned} &P_{y^{(n)}}\big(e^{-n(U-2\kappa)}\tau^{(n)}_{\lfloor e^{n(U-2\kappa)}\rfloor} \leq e^{-n\kappa}\big) \\ &\leq P_{y^{(n)}}\big(e^{-n(U-2\kappa)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} (\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1 \leq e^{-n\kappa}\big) \\ &= P_{y^{(n)}}\big(e^{-n(U-2\kappa)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} \big((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1 - E_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1)\big) \\ &\leq e^{-n\kappa} - e^{-n(U-2\kappa)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} E_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1)\big) \\ &\leq P_{y^{(n)}}\big(e^{-n(U-2\kappa)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} \big((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1 - E_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1)\big) \\ &\leq e^{-n\kappa} - \frac{1}{2} \inf_{y' \in B_{3\eta}^{(n)} \setminus B_{2\eta}^{(n)}} E_{y'}(\tau_0^{(n)} \wedge 1)\big)\,. \end{aligned}$$

Let $T' > 0$ denote the infimum of time lengths it takes a solution of (1.1) to get from a point in $B_{3\eta} \setminus B_{2\eta}$ to a point in B_η. Since $Y^{(n)}$ started at point $y'^{(n)} \in B_{3\eta}^{(n)} \setminus B_{2\eta}^{(n)}$ such that $y'^{(n)} \to y'$ converges to the solution of (1.1) started at y' locally uniformly, $\tau_0^{(n)}$ is greater than $T'/2$ with great $P_{y'^{(n)}}$–probability, for n great enough. Therefore, $\liminf_{n\to\infty} E_{y'^{(n)}}(\tau_0^{(n)} \wedge 1) \geq (T'/2) \wedge 1$. Hence, $e^{-n\kappa} - (1/2)\inf_{y' \in B_{3\eta}^{(n)} \setminus B_{2\eta}^{(n)}} E_{y'}(\tau_0^{(n)} \wedge 1) < 0$, for all n great enough, so that, on recalling that the $\tau_i^{(n)} - \sigma_i^{(n)}$, for $i \in \mathbb{N}$, are independent, by Chebyshev's inequality,

$$P_{y^{(n)}}\Big(e^{-n(U-2\kappa)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)}\rfloor} \big((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1 - E_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1)\big)$$

$$\leq e^{-n\kappa} - \frac{1}{2} \inf_{y' \in B_{3\eta}^{(n)} \setminus B_{2\eta}^{(n)}} E_{y'}(\tau_0^{(n)} \wedge 1)\Big)$$

$$\leq P_{y^{(n)}}\Big(e^{-n(U-2\kappa)}\Big| \sum_{i=1}^{\lfloor e^{n(U-2\kappa)} \rfloor} \big((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1 - E_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1)\big)\Big|$$

$$\geq \frac{(T'/2) \wedge 1}{2} - e^{-n\kappa}\Big)$$

$$\leq \frac{e^{-2n(U-2\kappa)}}{(((T'/2) \wedge 1)/2 - e^{-n\kappa})^2} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)} \rfloor} \mathrm{Var}_{y^{(n)}}((\tau_i^{(n)} - \sigma_i^{(n)}) \wedge 1)$$

$$\leq \frac{e^{-n(U-2\kappa)}}{(((T'/2) \wedge 1)/2 - e^{-n\kappa})^2}\,.$$

Thus, the righthand side of (A.16) converges to 0 so that (A.2) has been proved.

References

[1] V.M. Alekseev, V.M. Tikhomirov, and S.V. Fomin (1979) *Optimal Control*. Nauka. In Russian, English translation: Springer Science+Business Media, New York, 1987.

[2] N. Antunes, C. Fricker, P. Robert, and D. Tibi (2008) Stochastic networks with multiple stable points. *Ann. Probab.* **36** (1), 255–278.

[3] S. Asmussen (2003) *Applied probability and queues*, volume 51 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition. Stochastic Modelling and Applied Probability.

[4] D. Cox, J. Little, and D. O'Shea (1992) *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer-Verlag, New York. An introduction to computational algebraic geometry and commutative algebra.

[5] S.N. Ethier and T.G. Kurtz (1986) *Markov processes. Characterization and convergence*. Wiley.

[6] M.I. Freidlin and A.D. Wentzell (1998) *Random Perturbations of Dynamical Systems*. Springer, 2nd edition.

[7] J. Jacod and A.N. Shiryaev (1987) *Limit Theorems for Stochastic Processes*. Springer.

[8] A. Puhalskii (2001) *Large Deviations and Idempotent Probability*. Chapman & Hall/CRC.

[9] A. Puhalskii (2021) Large deviation limits of invariant measures. Arxiv preprint `arxiv:2006.16456v2`.

[10] A. Shwartz and A. Weiss (1995) *Large deviations for performance analysis*. Stochastic Modeling Series. Chapman & Hall, London. Queues, communication, and computing, with an appendix by R.J. Vanderbei.

[11] A. Shwartz and A. Weiss (2005) Large deviations with diminishing rates. *Math. Oper. Res.* **30** (2), 281–310.

[12] D. Tibi (2010) Metastability in communication networks. Arxiv preprint `arxiv:1002.0796v2`.

[13] A.D. Wentzell (1990) *Limit theorems on large deviations for Markov stochastic processes*, volume 38 of *Mathematics and its Applications* (*Soviet Series*). Kluwer Academic Publishers Group, Dordrecht. Translated from the Russian.