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 $\alpha_k$, $1 \leq k \leq K$. 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 $X^{(n)}(i,t) = (X^{(n)}_{i,1}(t),...,X^{(n)}_{i,K}(t))$, whose $k$-th entry records the number of class $k$ customers at the node. The process $X^{(n)}(i) = (X^{(n)}(i,t),t \geq 0)$ 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 $\Theta$. Let $Y^{(n)}(t)$ represent the proportion of nodes with $\theta$ as the population vector, i.e.,

$$Y^{(n)}(t) = \frac{1}{n} \sum_{i=1}^n 1\{X^{(n)}(i,t) = \theta\}$$

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

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

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

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

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

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

(1.1)

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

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

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

Both $Y^{(n)}(t)$ and $y(t)$ are probability distributions on $\Theta$. 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 $\Theta$ be defined as

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

(1.2)

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

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

(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 trajectory 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 $y \rightarrow y + (f_{\theta + e_{k-1}} - f_{\theta})/n$, which correspond to departures of class $k$ customers from nodes with population vector $\theta$, occur at the rate $n y_\theta \theta_k \delta_k$, which, when divided by $n$, tends to 0 as $y_\theta \rightarrow 0$, where $f_\theta$ denotes the $\theta$-th vector of the canonical basis of $\mathbb{R}^\Theta$. The line of attack in this paper is to prove $C$–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 \( \lambda = (\lambda_\theta, \theta \in \Theta) \in \mathbb{R}^\Theta \),

\[
H(y, \lambda) = \sum_{k=1}^{K} \sum_{\theta \in \Theta^+_k} (e^{\lambda_{\theta+k} - \lambda_{\theta}} - 1)\alpha_k y_\theta + \sum_{k=1}^{K} \sum_{\theta \in \Theta^-_k} (e^{\lambda_{\theta-k} - \lambda_{\theta}} - 1)(\delta_k + \gamma_k \sum_{\theta' \in \Theta \setminus \Theta_k^+} y_{\theta'} \theta_k y_\theta)
\]

\[
- \sum_{k=1}^{K} \sum_{\theta' \in \Theta^-_k, \theta \in \Theta_k^+} (e^{\lambda_{\theta'-k} + \lambda_{\theta} - \lambda_{\theta'} - \lambda_{\theta}} - 1)\theta_k \gamma_k y_\theta y_{\theta'} \tag{2.1}
\]

and

\[
L(y, z) = \sup_{\lambda \in \mathbb{R}^\Theta} \left( \lambda \cdot z - H(y, \lambda) \right), \tag{2.2}
\]

where \( \Theta_k^\pm = \{ \theta \in \Theta : \theta \pm e_k \in \Theta \} \) and “ \( \cdot \) ” is used to denote an inner product. Let \( \mathcal{D}(\mathbb{R}_+, \mathbb{R}^\Theta) \) 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 \( y^{(n)} \in S^{(n)}_\Theta \), \( y \in S\), 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

\[
I^*_y(y) = \int_0^\infty L(y(s), \hat{y}(s)) \, ds,
\]

provided \( y = (y(t), t \geq 0) \) is an absolutely continuous function taking values in \( S\) with \( y(0) = y \), and \( I^*_y(y) = \infty \), otherwise.

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

\[
(1/n) \ln \mathbf{P}(Y^{(n)} \in W) \to - \inf_{y \in W} I^*_y(y),
\]
as \( n \to \infty \), where \( \text{int} \) and \( \text{cl} \) 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 \( \hat{y}(t) = \nabla_\lambda H(y(t), \lambda) \big|_{\lambda = 0} \). In addition, \( I^*_y(y') = 0 \) if and only if \( y' = y \), with \( y(0) = y \).

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 \( \theta \) result in jumps \( f_{\theta + \epsilon_k}/n - f_{\theta}/n \), departures of class \( k \) customers from nodes with population vector \( \theta \) produce jumps \( f_{\theta - \epsilon_k}/n - f_{\theta}/n \), whereas class \( k \) customer migrations from nodes \( \theta \) to nodes \( \theta' \) give rise to jumps \( f_{\theta + \epsilon_k}/n - f_{\theta'}/n + f_{\theta - \epsilon_k}/n - f_{\theta'}/n \). 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 \( \Gamma \) 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,

\[
\mu^{(n)}([0, t], \{ f_{\theta + \epsilon_k}/n - f_{\theta}/n \}) = \int_0^t \sum_{i=1}^n \mathbf{1}_{\{ X^{(n)}(s-) = \theta \}} dN^{(n)}_{i,k}(s),
\]

\[
\mu^{(n)}([0, t], \{ f_{\theta - \epsilon_k}/n - f_{\theta}/n \}) = \int_0^t \sum_{i=1}^n \mathbf{1}_{\{ X^{(n)}(s-) = \theta \}} \sum_{j=1}^\infty \mathbf{1}_{\{ j \leq X^{(n)}_{i,k}(s-) \}} dL^{(n)}_{i,k,j}(s),
\]

where \( X^{(n)}_{i,k}(s) \) is the population of class \( k \) customers and \( X^{(n)}(s-) \) is the lefthand limit of \( X^{(n)}(s) \) at \( s \).
\[ A.A. \text{ Puhalskii} \]

\[ + \int_0^t \sum_{\theta' \in \Theta_0^+} \sum_{i=1}^n \sum_{i' = 1}^n 1_{(i' \neq i)} 1_{\{X_i^{(n)}(s-) = \theta\}} 1_{\{X_{i'}^{(n)}(s-) = \theta'\}} \]

\[ \sum_{j=1}^\infty 1_{\{j \leq X_{i,k}^{(n)}(s-)\}} 1_{\{e_{i,k,j}^{(n)}(s) = i\}} \text{d}R_{i,k,j}^{(n)}(s) \]

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

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

\[ \sum_{j=1}^\infty 1_{\{j \leq X_{i,k}^{(n)}(s-)\}} 1_{\{e_{i,k,j}^{(n)}(s) = i\}} \text{d}R_{i,k,j}^{(n)}(s) \]

where the \( N_{i,k}^{(n)} \), \( L_{i,k,j}^{(n)} \) and \( R_{i,k,j}^{(n)} \) are independent Poisson processes of respective rates \( \alpha_k \), \( \delta_k \) and \( \gamma_k \), 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 \( \{1, 2, \ldots, n\} \setminus \{i\} \), 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,

\[ \nu^{(n)}([0, t], \{f_{\theta' + e_{k}} \frac{n}{n} - f_{\theta} \frac{n}{n} + f_{\theta - e_{k}} \frac{n}{n} - f_{\theta} \frac{n}{n}\}) = n \int_0^t Y_{\theta}^{(n)}(s) \alpha_k \text{d}s \]

\[ \nu^{(n)}([0, t], \{f_{\theta - e_{k}} \frac{n}{n} - f_{\theta} \frac{n}{n}\}) = n \int_0^t Y_{\theta}^{(n)}(s) \theta k \gamma_k \text{d}s \]

\[ + \frac{n^2}{n - 1} \int_0^t Y_{\theta}^{(n)}(s) \sum_{\theta' \in \Theta_0^+} Y_{\theta'}^{(n)}(s) 1_{\{\theta' \in \Theta \setminus \Theta_0^+\}} \theta k \gamma_k \text{d}s \]

\[ \nu^{(n)}([0, t], \{f_{\theta - e_{k}} \frac{n}{n} - f_{\theta} \frac{n}{n}\}) = \frac{n^2}{n - 1} \int_0^t Y_{\theta}^{(n)}(s) Y_{\theta'}^{(n)}(s) \theta k \gamma_k \text{d}s \]
On the metastability

\[- \frac{n}{n-1} \int_0^t Y^{(n)}_{\theta}(s) 1_{\{y' = \theta \}} \theta_k \gamma_k \, ds.\]

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

\[G^{(n)}_t(\lambda) = t \int_0^t (\exp(\lambda u - 1) - 1) n \int_0^t Y^{(n)}_{\theta}(s) \alpha_k \, ds \]

\[+ \sum_{k=1}^K \sum_{\theta \in \Theta_+^k} \left( e^{(\lambda \theta - \lambda \theta) / n} - 1 \right) \left( n \int_0^t Y^{(n)}_{\theta}(s) \theta_k \delta_k \, ds \right) \]

\[+ \frac{n^2}{n-1} \int_0^t Y^{(n)}_{\theta}(s) Y^{(n)}_{\theta'}(s) 1_{\{\theta' \in \Theta \}} \theta_k \gamma_k \, ds \]

\[+ \sum_{k=1}^K \sum_{\theta \in \Theta_+^k} \left( e^{(\lambda \theta - \lambda \theta) / n} - 1 \right) \left( n \int_0^t Y^{(n)}_{\theta}(s) 1_{\{\theta' = \theta \}} \theta_k \gamma_k \, ds \right) \]

\[+ \frac{n^2}{n-1} \int_0^t Y^{(n)}_{\theta}(s) Y^{(n)}_{\theta'}(s) \theta_k \gamma_k \, ds - \frac{n}{n-1} \int_0^t Y^{(n)}_{\theta}(s) 1_{\{\theta' = \theta \}} \theta_k \gamma_k \, ds.\]

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 \(E\)) on p. 357 in Puhalskii [8] need be checked with \(E^{(n)}_t(\lambda) = e^{G^{(n)}_t(\lambda)}\), see (4.1.15) on p. 293 in Puhalskii [8]. Recalling that \(Y^{(n)}(t)\) takes values in [0,1] implies that the condition in question holds with

\[G_t(\lambda; y) = t \int_0^t H(y(s), \lambda) \, ds.\]

By Theorem 5.1.5 on p.357 in Puhalskii [8], the sequence \(Y^{(n)}\) is \(C\)-exponentially tight in \(\mathbb{D}(\mathbb{R}_+, \mathbb{R}^\Theta)\) and its every LD limit point solves maxingale problem \((y, G)\).
Let deviability $\Pi_y = (\Pi_y(W), W \subset \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d))$ represent an LD limit point of $Y^{(n)}$ (recall that $y^{(n)} \to y$), i.e., $\Pi_y(W) \in [0, 1]$, $\Pi_y(\emptyset) = 0$, $\Pi_y(\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)) = 1$, $\Pi_y(W) = \sup_{y \in W} \Pi_y(y)$, sets $\{y \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d) : \Pi_y(y) \geq \beta\}$ are compact for $\beta \in (0, 1]$, where $\Pi_y(y) = \Pi_y(\{y\})$, and the sequence of the distributions of $Y^{(n)}$ obeys a subsequential LDP with deviation function $-\ln \Pi_y(y)$, see Puhalskii [8, 9]. Then, $\Pi_y(y) = 0$ unless $y \in C(\mathbb{R}_+, \mathbb{R}^d)$, $y(0) = y$, and $\exp(\lambda \cdot (y(t) - y) - G_\lambda(y))$ is a local exponential maxingale in $C(\mathbb{R}_+, \mathbb{R}^d)$, as defined in Puhalskii [8], where $C(\mathbb{R}_+, \mathbb{R}^d)$ denotes the subset of continuous functions of $\mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$ with the subspace topology which is the topology of locally uniform convergence.

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

$$\Pi_y(y) \leq \Pi'_y(y),$$

(2.4)

where

$$\Pi'_y(y) = e^{-I^*_y(y)},$$

(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 $\Pi'_y(y) > 0$. It is immediate that $\Pi'_y(y) = 0$ unless $y \in S_\Theta$ so that when proving the equality it may and will be assumed that $y(t) = (y_\theta(t), \theta \in \Theta) \in S_\Theta$, for all $t$. The equality in (2.4) is proved, at first, for the case where $y$ stays away from the boundary of $S_\Theta$ so that $y_\theta(t) > 0$, for all $\theta$ and $t$, see Lemma 2.3 below. Furthermore, if $y_\theta(s) > 0$ for all $s \in [0, t]$ and one defines, in analogy with pp.210, 212 in Puhalskii [8],

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

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

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

(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 $L(y^\varepsilon(s), \dot{y}^\varepsilon(s))$ is obtained through the use of a nontrivial bound on the optimisers in (2.2), see Lemma 2.2. Upper semicontinuity of $\Pi_y(y)$ in $(y, y)$ 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 \( \Pi_y(p_t^{-1}(p,y)) = \Pi_y(y) \), and a Markov property of \( \Pi_y \) is used in order to obtain the needed equality \( \Pi_y(y) = \Pi_y(y) \).

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

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

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

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

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

**Proof.** For \( \hat{\lambda} = (\hat{\lambda}_{\theta, k}) \in \prod_{k \in \{1, 2, \ldots, K\}} \mathbb{R}^{\Theta_k} \), define

\[
\hat{H}(y, \hat{\lambda}) = \sum_{k=1}^{K} \sum_{\theta \in \Theta_k^-} (e^{\hat{\lambda}_{\theta, k}} - 1) \alpha_k y_\theta + \sum_{k=1}^{K} \sum_{\theta \in \Theta_k^+} (e^{-\hat{\lambda}_{\theta, -e_k}} - 1) (\delta_k + \gamma_k \sum_{\theta' \in \Theta \setminus \Theta_k^+} y_{\theta'} \theta_k y_\theta) + \sum_{k=1}^{K} \sum_{\theta \in \Theta_k^-} (e^{\hat{\lambda}_{\theta', k}} - \hat{\lambda}_{\theta', -e_k} - 1) \theta_k \gamma_k y_\theta y_{\theta'}.
\]
so that
\[ L(y, z) = \sup_{\lambda = (\lambda_{0}) \in \mathbb{R}^{\theta}} (\lambda \cdot z - \hat{H}(y, \lambda)) = \sup_{\lambda = (\lambda_{0}) \in \mathbb{R}^{\theta^{+}}}, \lambda = (\lambda_{0}, k) \in \prod_{k=1,2,\ldots,K} \mathbb{R}^{y_{k}^{+}}: \lambda_{0} + e_{k} - \lambda_{0} - \lambda_{0,k} = 0 \] (2.7)

Define a Lagrange function, with \( r_{\theta,k} \in \mathbb{R} \) and \( r = (r_{\theta,k}), \)
\[ L(\lambda, \hat{\lambda}, r, y, z) = \sum_{\theta \in \Theta} \lambda_{0} z_{\theta} - \hat{H}(y, \hat{\lambda}) + \sum_{\theta \in \Theta^{+}} r_{\theta,k} (\lambda_{0} + e_{k} - \lambda_{0} - \hat{\lambda}_{0,k}). \]

The optimality conditions in (2.7) that \( \partial_{\lambda_{0,k}} L(\lambda, \hat{\lambda}, r, y, z) = 0 \) and \( \partial_{\lambda_{0}} L(\lambda, \hat{\lambda}, r, y, z) = 0 \), 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
\[ -e^{-\hat{\lambda}_{0,k}} (\alpha_{k} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} (\theta_{k} + 1) y_{\theta' + e_{k}} y_{\theta} = 0, \text{ for } \theta \in \Theta^{+}_{k} \text{ with } \theta_{k} \geq 1, \] (2.9a)
\[ -e^{-\hat{\lambda}_{0,k}} (\theta_{k} + 1) (\delta_{k} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} y_{\theta'} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} y_{\theta' + e_{k}} y_{\theta} = r_{\theta,k} \] (2.9b)

Summing in (2.8) yields
\[ -\sum_{\theta \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} \alpha_{k} y_{\theta} + \sum_{\theta \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} (\delta_{k} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} y_{\theta'}) (\theta_{k} + 1) y_{\theta' + e_{k}} = \sum_{\theta \in \Theta^{+}} r_{\theta,k}. \] (2.10)

Solving for \( \sum_{\theta \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} y_{\theta} \) and substituting in (2.8) imply, after some algebra, that, for \( \theta \in \Theta^{+}_{k} \),
\[ \alpha_{k} e^{-\hat{\lambda}_{0,k}} y_{\theta} = -e^{-\hat{\lambda}_{0,k}} y_{\theta' + e_{k}} (\theta_{k} + 1) \]
\[ \times (\delta_{k} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} y_{\theta'}) = \frac{\gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} r_{\theta',k}}{\alpha_{k} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} (\theta_{k} + 1) y_{\theta' + e_{k}}} \]
\[ = \frac{-\alpha_{k} r_{\theta,k}}{\alpha_{k} + \gamma_{k} \sum_{\theta' \in \Theta^{+}_{k}} e^{-\hat{\lambda}_{0,k}} (\theta_{k} + 1) y_{\theta' + e_{k}}} \].
Therefore, if ˆλθ,k > 0, then
\[ \alpha_k e^{\hat{\lambda}_{\theta,k} y_0} \leq y_{\theta+e_k} (\theta_k + 1) \]
\[ \times \left( \delta_k + \gamma_k \sum_{\theta' \in \Theta^+} y_{\theta'} + \frac{\gamma_k |\sum_{\theta' \in \Theta^+} r_{\theta',k}|}{\alpha_k + \gamma_k \sum_{\theta' \in \Theta^+} e^{-\hat{\lambda}_{\theta',k}} (\theta'_k + 1)y_{\theta'+e_k}} \right) \]
\[ - \frac{\alpha_k r_{\theta,k}}{\alpha_k + \gamma_k \sum_{\theta' \in \Theta^+} e^{-\hat{\lambda}_{\theta',k}} (\theta'_k + 1)y_{\theta'+e_k}} \]
\[ \leq y_{\theta+e_k} (\theta_k + 1) \left( \delta_k + \gamma_k + \frac{\gamma_k |\sum_{\theta' \in \Theta^+} r_{\theta',k}|}{\alpha_k} \right) + |r_{\theta,k}|, \]

which implies that, no matter the sign of ˆλθ,k,
\[ \alpha_k e^{\hat{\lambda}_{\theta,k} y_0} \leq \alpha_k y_0 + y_{\theta+e_k} (\theta_k + 1) \left( \delta_k + \gamma_k + \frac{\gamma_k |\sum_{\theta' \in \Theta^+} r_{\theta',k}|}{\alpha_k} \right) + |r_{\theta,k}|. \quad \text{(2.11)} \]

By (2.10) and (2.11),
\[ \sum_{\theta \in \Theta} e^{-\hat{\lambda}_{\theta,k}} \sigma_k (\theta_k + 1)y_{\theta+k} \]
\[ \leq \sum_{\theta \in \Theta} (\alpha_k y_0 + y_{\theta+e_k} (\theta_k + 1) \left( \delta_k + \gamma_k + \frac{\gamma_k |\sum_{\theta' \in \Theta^+} r_{\theta',k}|}{\alpha_k} \right) + |r_{\theta,k}|) \]
\[ + \sum_{\theta \in \Theta} r_{\theta,k}. \quad \text{(2.12)} \]

Solving (2.9a) and (2.9b) recursively yields
\[ r_{\theta,k} = \sum_{i=0}^{\theta_k} z_{\theta-i e_k}. \quad \text{(2.13)} \]

As a consequence of (2.11), (2.12), and (2.13), for some C_1 > 0 and C_2 > 0,
\[ e^{\hat{\lambda}_{\theta,k} y_0} + e^{-\hat{\lambda}_{\theta,k} y_{\theta+e_k}} \leq C_1 + C_2 \sum_{\theta' \in \Theta} |z_{\theta'}|. \quad \text{(2.14)} \]

By (2.8), in analogy with (2.11), for \( \theta \in \Theta^+_k \),
\[ e^{\hat{\lambda}_{\theta,k} (\alpha_k + \gamma_k \sum_{\theta' \in \Theta^+_k} e^{-\hat{\lambda}_{\theta',k}} (\theta'_k + 1)y_{\theta'+e_k}) y_0} \]
\[ \leq (\theta_k + 1) \left( \delta_k + \gamma_k \sum_{\theta' \in \Theta^+_k} y_{\theta'} + \gamma_k \sum_{\theta' \in \Theta^+_k} e^{\hat{\lambda}_{\theta',k} y_{\theta'+e_k}} \right) + |r_{\theta,k}| \]
By (2.13) and (2.14), there exist $C''_1 > 0$ and $C''_2 > 0$ such that, for $\theta \in \Theta^+_k$ and $\theta' \in \Theta^-_k$,

$$e^{\lambda_{\theta,k} - \lambda_{\theta' - c_k,k}y_\theta} \leq C''_1 + C''_2 \sum_{\theta'' \in \Theta} |\bar{\gamma}_{\theta''}|,$$

which concludes the proof on recalling that $\bar{\lambda}_{\theta,k} = \lambda_{\theta + c_k} - \lambda_0$ and that $\bar{\lambda}_{\theta' - c_k,k} = \lambda_{\theta'} - \lambda_{\theta' - c_k}$.

Denote $\Pi_{\nu,\lambda}(y) = \Pi_{\nu}(p_{\nu}^{-1}(p_{\nu}y))$.

**Lemma 2.3.** Let $y = (y(t), t \geq 0)$ be an absolutely continuous function taking values in $S_{|\lambda|}$ with $y(0) = y$. If $y$ is locally bounded away from zero entrywise, then $\Pi_{\theta,\lambda}(y) = \Pi_{\theta,\lambda}(y)$, for all $t$, and $\Pi_{\nu}(y) = \Pi_{\nu}(y)$.

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

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

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

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

Calculations, using (2.1), yield

$$\dot{y}_\theta(s) = (\alpha_k + \sum_{\theta'' \in \Theta^-_k} \theta'' \gamma_k e^{\lambda_{\theta' - c_k}(s) - \lambda_{\theta''}(s)}y_{\theta''}(s))e^{\lambda_\theta(s) - \lambda_{\theta - c_k}(s)}y_{\theta - c_k}(s)1_{\theta \in \Theta^+_k}$$

$$+ \left(\delta_k + \gamma_k \sum_{\theta' \in \Theta^+_k} y_{\theta'}(s) + \gamma_k \sum_{\theta'' \in \Theta^-_k} e^{\lambda_{\theta' + c_k}(s) - \lambda_{\theta''}(s)}y_{\theta''}(s)\right)$$

$$\times e^{\lambda_\theta(s) - \lambda_{\theta + c_k}(s)}(\theta_k + 1)y_{\theta + c_k}(s)1_{\theta \in \Theta^+_k}$$

$$- \left((\alpha_k + \sum_{\theta'' \in \Theta^-_k} \theta'' \gamma_k e^{\lambda_{\theta' - c_k}(s) - \lambda_{\theta''}(s)}y_{\theta''}(s))e^{\lambda_{\theta - c_k}(s) - \lambda_\theta(s)}1_{\theta \in \Theta^+_k}\right)$$

$$+ \left(\delta_k + \gamma_k \sum_{\theta' \in \Theta^+_k} y_{\theta'}(s) + \gamma_k \sum_{\theta'' \in \Theta^-_k} \right)$$

$$\times e^{\lambda_{\theta' + c_k}(s) - \lambda_{\theta''}(s)}y_{\theta''}(s)e^{\lambda_{\theta - c_k}(s) - \lambda_\theta(s)}\theta_k1_{\theta \in \Theta^+_k})y_{\theta}(s).$$

Since the $y_\theta(s)$ are locally bounded away from zero, Lemma 2.2 implies that the exponentials on the latter righthand side are locally integrable functions of
Lemma 2.4. Let \( y \) be an absolutely continuous function with values in \( S_{|\Theta|} \). Let \( \theta^* \) represent a point of the maximum of \( y_\theta(0) \) so that \( y_\theta^* (0) = \max_{\theta \in \Theta} y_\theta(0) \). For \( \varepsilon \in (0, 1/(3|\Theta|^2)) \), let \( y_\varepsilon^\theta(s) = y_\theta(s) + \varepsilon \) unless \( \theta = \theta^* \) and let \( y_\varepsilon^\theta(s) = 1 - \sum_{\theta \neq \theta^*} y_\theta^\theta(s) \). Then, for \( t \) such that

\[
\sum_{\theta \in \Theta} \int_0^t |y_\theta(s)| \, ds \leq 1/(3|\Theta|),
\]

\( y_\varepsilon^\theta(s) \in S_{|\Theta|} \) on \([0, t]\) and

\[
\lim_{\varepsilon \to 0} \int_0^t L(y_\varepsilon^\theta(s), y_\varepsilon^\theta(s)) \, ds = \int_0^t L(y(s), y(s)) \, ds.
\] (2.16)

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

\[
H(y_\varepsilon^\theta(s), \lambda) \geq \sum_{k=1}^K \left( \sum_{\theta \in \Theta_k^+ \setminus \{\theta^*\}} e^{\lambda_{\theta^* + k} - \lambda_\theta} \alpha_k y_\theta(s) - |\Theta| \varepsilon e^{\lambda_{\theta^* + k} - \lambda_\theta^*} \alpha_k \right)
\]

\[
- \sum_{k=1}^K \sum_{\theta \in \Theta_k^+ \setminus \{\theta^*\}} \alpha_k (y_\theta(s) + \varepsilon)
\]

\[
+ \sum_{k=1}^K \left( \sum_{\theta \in \Theta_k^+ \setminus \{\theta^*\}} e^{\lambda_{\theta^*} - \lambda_\theta} (\delta_k + \gamma_k (\sum_{\theta' \in \Theta \setminus \Theta_k^+} y_{\theta'}(s) - \varepsilon |\Theta|)) \theta_k y_\theta(s)
\]

\[
- e^{\lambda_{\theta^*} - \lambda_{\theta^*}} (\delta_k + \gamma_k) \theta_k^* |\Theta| \varepsilon
\]

\[
- \sum_{k=1}^K \sum_{\theta \in \Theta_k^+ \setminus \{\theta^*\}} (\delta_k + \gamma_k \sum_{\theta' \in \Theta \setminus \Theta_k^+} y_{\theta'}(s) + \gamma_k |\Theta|) \theta_k (y_\theta(s) + \varepsilon)
\]

\[
+ \sum_{k=1}^K \gamma_k \left( \sum_{\theta' \in \Theta_k^+ \setminus \{\theta^*\}} e^{\lambda_{\theta'} + k - \lambda_{\theta^*}} y_{\theta'}(s) - e^{\lambda_{\theta^* + k} - \lambda_{\theta^*}} |\Theta| \varepsilon \right)
\]
\[
\times \left( \sum_{\theta \in \Theta_k^-} e^{\lambda_{s_{-k}} - \lambda_s \theta_k y_\theta(s)} - e^{\lambda_{s_{-k}} + \lambda_s \theta_k^*} \theta_k^* |\Theta| \right)
\]
\[
- \sum_{k=1}^{K} \sum_{\theta \in \Theta_k^-} \theta_k \gamma_k (y_\theta(s) + \varepsilon)(y_\theta(s) + \varepsilon)
\]
\[
\geq H(y(s), \lambda) - |\Theta| \varepsilon \sum_{\theta \in \Theta_k^*} e^{\lambda_{s_{-k}} + \lambda_s \theta_k^*} \theta_k^* y_\theta(s)
\]
\[
+ \sum_{k=1}^{K} \sum_{\theta \in \Theta_k^*} e^{\lambda_{s_{-k}} - \lambda_s \theta_k y_\theta(s)}
\]
\[
+ \sum_{k=1}^{K} \gamma_k e^{\lambda_{s_{-k}} - \lambda_s} \sum_{\theta \in \Theta_k^-} e^{\lambda_{s_{-k}} - \lambda_s \theta_k y_\theta(s)}
\]
\[
\geq H(y(s), \lambda) - |\Theta| \varepsilon \sum_{\theta \in \Theta_k^*} e^{\lambda_{s_{-k}} + \lambda_s \theta_k^*} \theta_k^* y_\theta(s)
\]
\[
+ \sum_{k=1}^{K} \gamma_k e^{\lambda_{s_{-k}} - \lambda_s} \sum_{\theta \in \Theta_k^*} e^{\lambda_{s_{-k}} - \lambda_s \theta_k^*} y_\theta(s)
\]
\[
\geq H(y(s), \lambda) - |\Theta| \varepsilon \sum_{\theta \in \Theta_k^*} e^{\lambda_{s_{-k}} + \lambda_s \theta_k^*} \theta_k^* y_\theta(s)
\]
\[
+ \sum_{k=1}^{K} \gamma_k e^{\lambda_{s_{-k}} - \lambda_s} \sum_{\theta \in \Theta_k^-} e^{\lambda_{s_{-k}} + \lambda_s \theta_k y_\theta(s)}
\]

and \(M > 0\) depends neither on \(\lambda\) nor on \(y(s)\).

Therefore,

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

Since \(y^\varepsilon_\theta(s)\) is locally bounded away from zero on \([0, t]\) uniformly in \(\varepsilon\), Lemma 2.2 implies that

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

Thus,

\[
\limsup_{\varepsilon \to 0} \int_0^t L(y^\varepsilon(s), \dot{y}^\varepsilon(s)) \, ds \leq \int_0^t L(y(s), \dot{y}(s)) \, ds.
\]  \hspace{1cm} (2.17)
On the other hand, by (2.1) and (2.2), a.e.,

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

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

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

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

$$\Pi_y(y) = \Pi_y(t) \Pi_y(\vartheta_t y),$$

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

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

$$E_{y}^{(\alpha)}(f(Y^{(\alpha)})) = E_{y}^{(\alpha)}(E_{y}^{(\alpha)}((f((p_t Y^{(\alpha)}, \vartheta_t Y^{(\alpha)})))^n|F^{(\alpha)}(t)))$$

$$= E_{y}^{(\alpha)}(E_{Y^{(\alpha)}(t)}(f(u, Y^{(\alpha)}))^{n}|u = p_t Y^{(\alpha)}). \quad (2.18)$$

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

$$(E_{y}^{(\alpha)}(f(p_t u_n Y^{(\alpha)}))^{1/n})^{1/n} \to \sup_y f(p_t y, \tilde{y}) \Pi_y(\tilde{y}).$$

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

$$g_n(y^{(\alpha)}) \to \sup_y f(p_t y, \tilde{y}) \Pi_y(t)(\tilde{y}).$$

Therefore, accounting for (2.18),

$$(E_{y}^{(\alpha)}(f(Y^{(\alpha)}))^{1/n})^{1/n} = (E_{y}^{(\alpha)}(g_n(Y^{(\alpha)}))^{1/n})^{1/n} \to \sup_y \sup_{\tilde{y}} f(p_t y, \tilde{y}) \Pi_y(t)(\tilde{y}) \Pi_y(\tilde{y})$$

$$= \sup_y \sup_{\tilde{y}} f(p_t y, \tilde{y}) \Pi_y(t)(\tilde{y}) \Pi_y(p_t^{-1}(p_t y))$$

$$= \sup_y \sup_{\tilde{y}, \tilde{y}'} f(p_t y, \vartheta_t y') \Pi_y(t)(\vartheta_t y') \Pi_y(p_t^{-1}(p_t y)).$$
Lemma 2.6. The function $\Pi_y(y)$ is upper semicontinuous in $(y,y)$.

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 $I^\gamma$. Then the distributions of the pairs $(Y^{(n)}(0),Y^n)$ satisfy a subsequential LDP with $I^\gamma(y) - \ln \Pi_y(y)$. Since the latter quantity is lower semicontinuous in $(y,y)$ and $I^\gamma(y)$ is continuous in $y$, $\Pi_y(y)$ is upper semicontinuous in $(y,y)$.

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

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

(2.19)

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

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

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

\hfill \Box

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 $\mathbb{S}^{(n)}$, see, e.g., Asmussen [3], which is denoted by $\mu^{(n)}$. It is convenient to extend $\mu^{(n)}$ to the whole of $\mathbb{S}$ by letting $\mu^{(n)}(\mathbb{S}) \setminus
On the metastability 381

Given \( \rho \), \( \nu \) on the vertex set \( \rho \), many solutions satisfy an LDP with |

Let the simplex \( S_\Theta \) be endowed with the subspace topology.

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

\[
J(y) = \inf_{\rho \in A^*} (J(\nu(\rho)) + \Phi(\nu(\rho), y)).
\]

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

1. if \( y^{(n)} \rightarrow y \), then the distributions of \( Y^{(n)} \) satisfy an LDP with \( I^*_y \),
2. the function \( I^*_y(y) \) is lower semicontinuous in \( (y, y') \) and the set \( \cup_{y \in S_{[\Theta]}} \{ y \in C(\mathbb{R}^+, S_{[\Theta]}) : I^*_y(y) \leq \eta \} \) is compact, for all \( \eta \geq 0 \),
3. for all \( y \in C(\mathbb{R}^+, S_{[\Theta]}) \),

\[
I^*_y(y) = \inf_{y' \in p_2^{-1}(y, y')} I^*_y(y') + I^*_{y(\Theta)}(\theta_s y),
\]
4. (a) if \( I^*_y(y) = 0 \), then \( \inf_{t \geq 0} d(y(t), A) = 0 \), where \( d \) is a metric on \( S_{[\rho]} \),
(b) if \( y(t) = \nu(\rho) \), for all \( t \geq 0 \), then \( I^*_y(\nu(\rho)) = 0 \), where \( \rho \in A \),
(c) for any \( \rho, \bar{\rho} \in A \), there exists \( t > 0 \) such that \( \Phi_s(\nu(\rho), \nu(\bar{\rho})) < \infty \),
(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 \( y \in S_{[\rho]} \) 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 \( s \leq t \), \( \lambda \in \mathbb{R}^\Omega \), and \( \varepsilon > 0 \),

\[
\lambda \cdot (y(t) - y(s)) \leq \varepsilon \int_s^t L(\bar{y}(u), \tilde{y}(u)) \, du + \varepsilon \int_s^t H(\bar{y}(u), \tilde{y}(u)) \, du.
\]

As \( t - s \rightarrow 0 \), with \( s \) and \( t \) being bounded, the second term on the right-hand side goes to 0 uniformly over \( y \) and over \( \lambda \) from a bounded set. The first term is bounded above by \( \varepsilon \int_0^\infty L(\bar{y}(u), \tilde{y}(u)) \, du \leq \varepsilon \eta \), 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), \( \inf_{y' \in \mathcal{P}_{[\rho]}(y)} I^*_y(y') = \int_0^\infty L(\bar{y}(t), \tilde{y}(t)) \, dt \) and \( I^*_y(\partial_s y) = \int_0^\infty \int_0^\infty L(\partial_s y(t), d/dt(\partial_s y(t))) \, dt = \int_0^\infty L(\bar{y}(t), \tilde{y}(t)) \, dt \).

As for part 4, Proposition 4 in Antunes et al. [2] implies that if \( y \) satisfies (1.1), then \( y(t) \) converges, as \( t \rightarrow \infty \), 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 \( y(s) = (1 - s \wedge 1)\nu(\rho) + s \wedge 1 \nu(\bar{\rho}) \).

Part 4(d) is addressed next. Given \( \rho \) such that \( 0 < d(y, \nu(\rho)) < \eta \), one lets \( y(t) = y + t(\nu(\rho) - y)/d(\nu(\rho), y) \). Then, \( y(0) = y \), \( y(d(\nu(\rho), y)) = \nu(\rho) \) and

\[
\Phi_d(\nu(\rho), y)(\nu(\rho), y) \leq \int_0^{d(\nu(\rho), y)} L(y(t), \tilde{y}(t)) \, dt. \tag{3.3}
\]

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

\[
H(y(t), \lambda) \geq \min_{k} (\alpha_k + \delta_ky_k) \sum_{k=1}^K \left( \sum_{\theta \in \Theta_k^*} e^{\lambda_\theta \cdot \alpha_k(\delta_k - \lambda_\theta \cdot \gamma_k)} \right) \frac{\kappa}{d(\nu(\rho), y)} t
- \sum_{k=1}^K (\alpha_k + |\Theta| (\delta_k + \gamma_k)). \tag{3.4}
\]
Let \( \lambda_\tilde{\theta} = \min_{\theta \in \Theta} \lambda_\theta \) and \( \lambda_\hat{\theta} = \max_{\theta \in \Theta} \lambda_\theta \). Let \( \tilde{\theta} = \theta_0, \theta_1, \ldots, \theta_k = \hat{\theta} \) represent a path from \( \tilde{\theta} \) to \( \hat{\theta} \), as defined in Remark 2.3. By Jensen’s inequality,

\[
\sum_{k=1}^{K} \sum_{\theta \in \Theta_k} e^{\lambda_{\theta k} - \lambda_{\theta k - 1}} \geq \sum_{i=1}^{\ell} e^{\lambda_{\theta i} - \lambda_{\theta i - 1}} \geq \ell \epsilon \sum_{i=1}^{\ell} (\lambda_{\theta i} - \lambda_{\theta i - 1}) / \ell = \ell \epsilon (\lambda_{\hat{\theta}} - \lambda_{\tilde{\theta}}) / \ell.
\]

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

\[
\sum_{\theta \in \Theta} \lambda_\theta y_\theta(t) - H(y(t), \lambda) \leq \left| \Theta \right| (\lambda_\tilde{\theta} - \lambda_\hat{\theta}) \frac{\max_{\theta \in \Theta} |\nu_\theta(\rho) - y_\theta|}{d(\nu(\rho), y)}
\]

\[
- \min_{k} (\alpha_k, \delta_k) e^{(\lambda_\theta - \lambda_{\hat{\theta}})/\Theta_k} \frac{K}{d(\nu(\rho), y)} t
\]

\[
+ \sum_{k=1}^{K} (\alpha_k + |\Theta| (\delta_k + \gamma_k)).
\]

A similar inequality holds if all the \( \lambda_\theta \) in (3.4) are the same. Maximisation over \( \lambda_\tilde{\theta} - \lambda_\hat{\theta} \) 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 \( y(t) = \nu(\rho) + t(y - \nu(\rho))/d(\nu(\rho), y) \), notes that \( y_\theta(t) \geq (1 - t)\nu_\theta(\rho)/d(\nu(\rho), y) \) 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 \( \rho', \rho'' \), then \( \rho' \) 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 \( A \),

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

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 trajectoryal 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 \( S_\Theta \), 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 \( \tau^{(n)} = \inf\{t \geq 0 : Y^{(n)}(t) \notin D\} \). Let \( y^{(n)} \in D \cap S_\Theta^{(n)} \). If \( y^{(n)} \rightarrow y \in D \), as \( n \rightarrow \infty \), then

\[
P_{y^{(n)}}\left(\left\| \frac{1}{n} \ln \tau^{(n)} - U \right\| > \kappa \right) \rightarrow 0
\]

and

\[
\frac{1}{n} \ln E_{y^{(n)}}(\tau^{(n)})^{m} \rightarrow mU,
\]

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

\[
g(y) = \sum_{\theta \in \Theta} y_{\theta} \ln \left( \prod_{k=1}^{K} \theta_{k}^{\gamma_{k}y_{\theta}} \right) - \sum_{k=1}^{K} \frac{\delta_{k} + \gamma_{k}}{\gamma_{k}} (u \ln u - u) \bigg|_{u = \alpha_{k} + \gamma_{k} \sum_{\theta \in \Theta} \theta_{k} y_{\theta}} (\rho_{k} y_{\theta}) / (\delta_{k} + \gamma_{k})
\]

In the interior of \( S_\Theta \), see Antunes et al. [2],

\[
\nabla g(y) \cdot V(y) = \sum_{k=1}^{K} \theta_{k} y_{\theta} \ln \left( \prod_{k=1}^{K} \theta_{k}^{\gamma_{k}y_{\theta}} \right) - \frac{b_{k}(y)}{\delta_{k} + \gamma_{k}} \theta_{k} y_{\theta},
\]

where

\[
b_{k}(y) = \alpha_{k} + \gamma_{k} \sum_{\theta \in \Theta} \theta_{k} y_{\theta}.
\]

Hence, \( \nabla g(y) \cdot V(y) \leq 0 \) so that \( g(y(t)) \) 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 \( S_\Theta \), is zero at \( y \): \( dg_{\nu(y)}(y) = 0 \). 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

\[
\phi(\rho) = - \ln Z(\rho) + \sum_{k=1}^{K} \left( \frac{\gamma_{k} + \delta_{k}}{\gamma_{k}} \rho_{k} - \frac{\alpha_{k}}{\gamma_{k}} \ln \rho_{k} \right).
\]

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

\[
g(\nu(\rho)) = \phi(\rho) + \sum_{k=1}^{K} \frac{\alpha_{k}}{\gamma_{k}} (\ln \frac{\alpha_{k}}{\gamma_{k}} + \delta_{k} - 1).
\]
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 $(\rho_1, \rho_2)$ 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, $\theta_1$ takes values in the set $\{0, 1, \ldots, C\}$ and $\theta_2$ takes values in $\{0, 1\}$. Then, with $\rho = (\rho_1, \rho_2)$,

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

and

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

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

$$Z(\rho) = \frac{\gamma_1 \rho_1 (\rho_2 + \rho^C_2 / C!)}{\alpha_1 - \delta_1 \rho_1}$$

(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 $k = 2$, $\rho_2$ satisfies the quadratic equation

$$\gamma_1 (\gamma_2 + \delta_2) \rho_2^2 + (\gamma_1 (\gamma_2 + \delta_2) \frac{\rho_1^C_1}{C!} - \alpha_2 \gamma_1 - \gamma_2 (\frac{\alpha_1}{\rho_1} - \delta_1)) \rho_2 - \alpha_2 \gamma_1 \frac{\rho^C_1}{C!} = 0.$$

Hence,

$$\rho_2(\rho_1) = \frac{1}{2\gamma_1 (\gamma_2 + \delta_2)} \left( -\gamma_1 (\gamma_2 + \delta_2) \frac{\rho_1^C_1}{C!} + \alpha_2 \gamma_1 + \gamma_2 (\frac{\alpha_1}{\rho_1} - \delta_1) \right) + \sqrt{\left( \gamma_1 (\gamma_2 + \delta_2) \frac{\rho_1^C_1}{C!} - \alpha_2 \gamma_1 - \gamma_2 (\frac{\alpha_1}{\rho_1} - \delta_1) \right)^2 + 4 \gamma_1^2 (\gamma_2 + \delta_2) \alpha_2 \frac{\rho_1^C_1}{C!}}.$$

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

$$h(\rho) = 0,$$
where
\[
h(\rho) = \sum_{i=0}^{C} \frac{\rho_1^i}{i!} + \rho_2 - \frac{\gamma_1 \rho_1}{\alpha_1 - \delta_1 \rho_1} \left( \frac{\rho_1^C}{C!} + \rho_2 \right).\]  
(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 \( \delta_1 \) and (or) \( \delta_2 \).) For \( \rho_1 \) close to zero, the term \( \alpha_1/\rho_1 \) on the right of (3.7) dominates, \( \rho_2 \) decreases rapidly as \( \rho_1 \) 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 \(1.5794 \times 10^5\) 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: \( \rho^{(1)} \approx (.5966, .8293) \), \( \rho^{(2)} \approx (4.1786, 8.1115) \), and \( \rho^{(3)} \approx (13.72715, 8.9906) \). The second derivatives of \( \phi \) are
\[
\nabla^2 \phi(\rho^{(1)}) \approx \begin{pmatrix} 1.2633 & 0.016025 \\ 0.016025 & 0.1243 \end{pmatrix},
\]
\[
\nabla^2 \phi(\rho^{(3)}) \approx \begin{pmatrix} 0.015412 & 1.1085 \times 10^{-6} \\ 1.1085 \times 10^{-6} & 0.1113 \end{pmatrix},
\]
and
\[ \nabla^2 \phi(\rho^{(2)}) \approx \begin{pmatrix} -0.069679 & 0.012120 \\ 0.012120 & 0.1370 \end{pmatrix}, \]
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 stable equilibria so that the network process spends exponentially long periods

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

Figure 3. 3D mesh plot and contour plot of \(\phi(\rho_1, \rho_2)\)
of time in the neighbourhoods of those equilibria, while \( \nu(\rho^{(2)}) \) is an unstable equilibrium.

The expected number of class \( k \) customers being

\[
EQ_k = \rho_k (1 - \sum_{\theta : \sum_j A_{kj} \theta_j > 0} \nu(\rho))
\]

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

\[
EQ_1 = \rho_1 (1 - \frac{1}{Z(\rho)} \rho_1^C)
\]

and

\[
EQ_2 = \rho_2 (1 - \frac{1}{Z(\rho)} \left( \sum_{\theta_i=1}^C \rho_1^{\theta_i} \rho_2 \right)) = \frac{\rho_2}{Z(\rho)},
\]

respectively. For the stable equilibria, calculations yield \((EQ_1^{(1)}, EQ_2^{(1)}) \approx (.596, .828)\) 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 \( \delta_1 \) and \( \delta_2 \). When \( \delta_1 = \delta_2 = .1 \), only a downward trend is present, so, there is only one equilibrium.)

Also, calculations yield \( \phi(\rho^{(1)}) = -12.284, \phi(\rho^{(2)}) = -11.56, \phi(\rho^{(3)}) = -12.043 \) so that \( g(\nu(\rho^{(2)})) > g(\nu(\rho^{(1)})) \lor g(\nu(\rho^{(3)})) \). Thus, if \( y(0) = \nu(\rho^{(2)}) \) experiences a small displacement \( \Delta \) 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 \( y(t) \) 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),

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

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

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

where

\[
\Psi = (\Phi(\nu^{(2)}, \nu^{(1)}) + \Phi(\nu^{(3)}, \nu^{(2)})) \wedge (\Phi(\nu^{(2)}, \nu^{(1)}) + \Phi(\nu^{(3)}, \nu^{(1)})) \wedge (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(1)}, \nu^{(3)})) \wedge (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(1)}, \nu^{(1)})) \]
On the metastability

\( \land (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(3)}, \nu^{(2)})) \land (\Phi(\nu^{(1)}, \nu^{(3)}) + \Phi(\nu^{(2)}, \nu^{(3)})) \)

\( \land (\Phi(\nu^{(2)}, \nu^{(3)}) + \Phi(\nu^{(1)}, \nu^{(2)})) \land (\Phi(\nu^{(2)}, \nu^{(3)}) + \Phi(\nu^{(1)}, \nu^{(3)})) \)

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

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

\[ J(\nu^{(1)}) = \Phi(\nu^{(3)}, \nu^{(2)}) \land \Phi(\nu^{(3)}, \nu^{(1)}) - \Psi, \]

\[ J(\nu^{(2)}) = (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(3)}, \nu^{(1)})) \land (\Phi(\nu^{(1)}, \nu^{(2)}) + \Phi(\nu^{(3)}, \nu^{(2)})) \land (\Phi(\nu^{(3)}, \nu^{(2)}) + \Phi(\nu^{(1)}, \nu^{(3)})) - \Psi, \]

\[ J(\nu^{(3)}) = (\Phi(\nu^{(2)}, \nu^{(3)}) + \Phi(\nu^{(1)}, \nu^{(2)})) \land \Phi(\nu^{(1)}, \nu^{(3)}) - \Psi \]

and

\[ \Psi = \Phi(\nu^{(3)}, \nu^{(2)}) \land \Phi(\nu^{(3)}, \nu^{(1)}) \land (\Phi(\nu^{(2)}, \nu^{(3)}) + \Phi(\nu^{(1)}, \nu^{(2)})) \land \Phi(\nu^{(1)}, \nu^{(3)}). \]

It is being conjectured that, furthermore, \( \Phi(\nu^{(2)}, \nu^{(1)}) = \Phi(\nu^{(2)}, \nu^{(3)}) = 0 \) and that \( \Phi(\nu^{(3)}, \nu^{(1)}) \leq \Phi(\nu^{(3)}, \nu^{(1)}) \) and \( \Phi(\nu^{(1)}, \nu^{(2)}) \leq \Phi(\nu^{(1)}, \nu^{(2)}) \), 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

\[ \lim_{n \to \infty} \sup \frac{1}{n} \ln E_{g(n)}(\tau^{(n)})^m \leq mU \quad (A.1) \]

and

\[ \lim_{n \to \infty} P_{g(n)}(\frac{1}{n} \ln \tau^{(n)} \leq U - 3\kappa) = 0. \quad (A.2) \]

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

\[ \lim_{n \to \infty} P_{g(n)}(\frac{1}{n} \ln \tau^{(n)} \geq U + \kappa)^{1/n} \leq \lim_{n \to \infty}(E_{g(n)}(\ln \tau^{(n)})^{1/n} e^{-(U+\kappa)} \leq e^{-\kappa} \]

so that

\[ \lim_{n \to \infty} P_{g(n)}(\frac{1}{n} \ln \tau^{(n)} \geq U + \kappa) = 0. \]

By Jensen's inequality, for \( \epsilon > 0 \),

\[ \frac{1}{n} \ln E_{g(n)}(\tau^{(n)} \lor \epsilon)^m \geq \frac{m}{n} E_{g(n)}(\ln \tau^{(n)} \lor \ln \epsilon). \]
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,
$$
\ln E_{y^{(n)}} (\tau^{(n)} \vee \varepsilon)^m \leq \ln(E_{y^{(n)}} (\tau^{(n)})^m + \varepsilon^m) \\
\leq \ln 2 + \ln(E_{y^{(n)}} (\tau^{(n)})^m) \vee \ln \varepsilon^m.
$$

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 $y$ of (1.1) with $y'$ as an initial point to reach $\cl B_{\eta/2}$, where $y' \in \cl D$. The function $T(y')$ is upper semicontinuous. Since $D$ is bounded, so is $T(y')$. Let $T_1 = \max_{y' \in \cl D}T(y')$. Let $\sigma^{(n)}$ represent the first time when $Y^{(n)}$ reaches the closed ball $\cl B_{\eta}$. Let $y(t)$ solve (1.1) with $y(0) = y$. Since $y^{(n)}(t) \to y$ and $y^n(t) \to y(t)$ uniformly on $[0, T]$ in $P_{y^{(n)}}$-probability, it may be assumed that
$$
P_{y^{(n)}}(\sigma^{(n)} \leq T_1) \geq \frac{1}{2},
$$
provided $n$ is great enough.

Lemma 3.1 in Puhalskii [9] implies that $\Phi(\nu(\rho), y')$ is a continuous function of $y'$. Therefore,
$$
U = \inf_{y' \in \cl D} \Phi(\nu(\rho), y').
$$

Let $D_\eta$ represent the open $\eta$–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 $\eta$ 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 $y^{(n)}$ and $t^{(n)}$ such that $y^{(n)}(0) = y^{(n)}$, $y^{(n)}(t^{(n)}) \notin D_\eta$ and $I_{y^{(n)}}(y^{(n)}) \leq U_\eta + \beta$. (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 $y^{(n)}(t) \notin D_\eta$. Let $s^{(n)}$ be the last time $t$ such that $y^{(n)}(t) \in \cl B_{\eta}$. Then, for $t \in [s^{(n)}, t^{(n)}]$, the function $y^{(n)}(t)$ takes values in the closed set $\cl D_{\eta} \setminus B_{\eta}$. Let $T$ denote the maximal length of time it takes a solution of (1.1) with an initial point in $\cl D_{\eta} \setminus B_{\eta}$ to get to $\cl B_{\eta}$. If
$t^{(n)} - s^{(n)} \leq NT$, for some $N$, the proof is over. Otherwise, on denoting, for $W \subset \mathbb{C}(\mathbb{R}_+, \mathbb{R}^{(\theta)})$, $I^*(W) = \inf_{y \in W} I^*(y)$, by (2.3),

$$I^*_{y^{(n)}}(y^{(n)}) \geq \inf_{y' \in clD \setminus B_n} I^*_y(\vartheta_{s^{(n)}} y^{(n)})$$

$$= \inf_{y' \in clD \setminus B_n} \left( I^*_y(p_T^{-1}(p_T(\vartheta_{s^{(n)}} y^{(n)}))) + I^*_y(\vartheta_{s^{(n)}} y^{(n)}+T y^{(n)}) \right)$$

$$\geq \inf_{y' \in clD \setminus B_n} I^*_y(\vartheta_{s^{(n)}} y^{(n)})$$

$$+ \inf_{y' \in clD \setminus B_n} I^*_y(p_T^{-1}(p_T(\vartheta_{s^{(n)}} y^{(n)})))$$

$$+ I^*_y(\vartheta_{s^{(n)}} y^{(n)}+2T y^{(n)}))$$

$$\geq \inf_{y' \in clD \setminus B_n} I^*_y(\vartheta_{s^{(n)}} y^{(n)}+2T y^{(n)}))$$

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

$$I^*_{y^{(n)}}(y^{(n)}) \geq \sum_{m=0}^{N} \inf_{y' \in clD \setminus B_n} I^*_y(p_T^{-1}(p_T(\vartheta_{s^{(n)}} y^{(n)}+mTy^{(n)})))$$

$$\geq N \inf_{y' \in clD \setminus B_n} I^*_y(y'),$$

for all $t \in [0, T]$.

The latter infimum is positive because no solution of (1.1) belongs to the set $\{y': y'(t) \in clD \setminus B_n \text{ for all } t \in [0, T]\}$. 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 $y^{(n)}$ by replacing the piece of $y^{(n)}$ on $[0, s^{(n)}]$ with a straight line segment connecting $y^{(n)}$ and $y^{(n)}(s^{(n)})$. The modified trajectory is given by $\tilde{y}^{(n)}(t) = y^{(n)}(t)$ for $t \in [0, s^{(n)}]$, and $\tilde{y}^{(n)}(t) = y^{(n)}(t - \vartheta_{s^{(n)}} y^{(n)} + s^{(n)})$, for $t \geq |y^{(n)}(s^{(n)}) - y^{(n)}|$. The $\tilde{y}^{(n)}$ are thus bounded. With $\bar{t}^{(n)} = s^{(n)} + |y^{(n)}(s^{(n)}) - y^{(n)}|$, $\bar{t}^{(n)}$ is the smallest $t$ with $\tilde{y}^{(n)}(t) \notin D$. In addition, it is possible to choose $\eta$ small enough to ensure that $I^*_{y^{(n)}}(\tilde{y}^{(n)}) \leq I^*_{y^{(n)}}(y^{(n)}) + \beta$. 

On the metastability
Since \( y'(n)(t^{(n)}) \notin D_q \), \( Y(n)(t^{(n)}) \notin D \) provided \( |Y(n)(t^{(n)}) - y(n)(t^{(n)})| < \eta \) so that

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

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

\[
\liminf_{n \to \infty} \frac{1}{n} \ln P_{y'}(\tau^{(n)} \leq T_2) \geq \liminf_{n \to \infty} \frac{1}{n} \ln P_{y'}(\tau^{(n)} \leq t^{(n)}) \\
\geq \liminf_{n \to \infty} \frac{1}{n} \ln P_{y'}(\sup_{s \leq t^{(n)}} |Y(n)(s) - y(n)(s)| < \eta) \\
\geq - \inf\{I_{y',\hat{y}}: \sup_{s \leq \hat{t}} |y(s) - \hat{y}(s)| < \eta\} \\
\geq -I_{y',\hat{y}}(\hat{y}) \geq -I_{y'}(\hat{y}) \geq -U - 2\beta \geq -U - 3\beta.
\]

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

\[
P_{y'}(\tau^{(n)} \leq T_1 + T_2) \geq P_{y'}(\sigma^{(n)} \leq T_1) \inf_{y'' \in B_r^{(n)}} P_{y''}(\tau^{(n)} \leq T_2),
\]

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 \( \ell \in N \),

\[
P_{y'}(\tau^{(n)} > \ell T_3 | \tau^{(n)} > (\ell - 1)T_3) = P_{y'}(Y(n)(t) \in D, t \in [(\ell - 1)T_3, \ell T_3]) | Y(n)(t) \in D, t \in [0, (\ell - 1)T_3]) \]

\[
= P_{y'}(Y(n)(t) \in D, t \in [0, T_3]) | Y(n)(0) \in D) \]

\[
\leq \sup_{y'' \in D^{(n)}} P_{y''}(\tau^{(n)} > T_3), \tag{A.6}
\]

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),

\[
\sup_{y' \in D^{(n)}} P_{y'}(\tau^{(n)} > \ell T_3) \leq (1 - e^{n(-U - \beta)})^\ell.
\]

Therefore,

\[
E_{y^{(n)}}(\tau^{(n)})^m = m \int_0^\infty u^{m-1} P_{y^{(n)}}(\tau^{(n)} > u) \, du
\]

\[
\leq mT_3 \sum_{\ell=0}^{\infty} ((\ell + 1)T_3)^{m-1} P_{y^{(n)}}(\tau^{(n)} > \ell T_3)
\]

\[
\leq mT_3^m \sum_{\ell=0}^{\infty} (\ell + 1)^{m-1} (1 - e^{n(-U - \beta)})^\ell
\]

\[
\leq mT_3^m \left(1 - e^{n(-U - \beta)}\right)^{-2} \int_0^\infty u^{m-1} (1 - e^{n(-U - \beta)})^u \, du
\]

\[
= mT_3^m \left(1 - e^{n(-U - \beta)}\right)^{-2} (\ln(1 - e^{n(-U - \beta)}))^m
\]

\[
\leq mT_3^m \left(1 - e^{n(-U - \beta)}\right)^{-2} e^{mn(U + \beta)}
\]

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

\[
P_{y^{(n)}}(\tau^{(n)} < e^{n(U - 3\kappa)})
\]

\[
= \sum_{i=0}^{\infty} P_{y^{(n)}}(\sigma_i^{(n)} \leq \tau^{(n)} < \tau_i^{(n)} \wedge \tau^{(n)} \leq e^{n(U - 3\kappa)})
\]

\[
\leq \sum_{i=0}^{[e^{n(U - 2\kappa)}]} P_{y^{(n)}}(\sigma_i^{(n)} \leq \tau^{(n)} < \tau_i^{(n)}) + P_{y^{(n)}}(\tau^{(n)} \leq e^{n(U - 3\kappa)})
\]

\[
= P_{y^{(n)}}(\tau^{(n)} < \tau_0^{(n)})
\]

\[
+ \sum_{i=1}^{[e^{n(U - 2\kappa)}]} E_{y^{(n)}}(1_{\{\sigma_i^{(n)} \leq \tau^{(n)}\}} P_{y^{(n)}}(\tau^{(n)} < \tau_i^{(n)} | \mathcal{F}^{(n)}(\sigma_i^{(n)})))
\]
\[ + P_{\hat{y}^{(n)}} \left( \frac{\tau_{\epsilon_0}^{(n)}}{e^{n(t - 2\eta)}} \right) < e^{-n\chi}. \]  

(A.7)

Since \( \tau^{(n)} = \inf \{ t \geq \sigma_{\epsilon_1}^{(n)} : Y^{(n)}(t) \notin D \} \) and \( \tau_{\epsilon_1}^{(n)} = \inf \{ t \geq \sigma_{\epsilon_1}^{(n)} : Y^{(n)}(t) \in B_{\eta} \} \) on the event \( \{ \sigma_{\epsilon_1}^{(n)} \leq \tau^{(n)} \} \), by the strong Markov property, for \( i \in \mathbb{N}, \ell \in \mathbb{N}, \) and \( n \) great enough,

\[
1_{\{ \sigma_{\epsilon_1}^{(n)} \leq \tau^{(n)} \}} P_{\hat{y}^{(n)}}(\tau^{(n)} < \tau_0^{(n)} | \mathcal{F}^{(n)}(\sigma_{\epsilon_1}^{(n)})) = 1_{\{ \sigma_{\epsilon_1}^{(n)} \leq \tau^{(n)} \}} P_{Y^{(n)}(\sigma_{\epsilon_1}^{(n)})}(\tau^{(n)} < \tau_0^{(n)}) \\
\leq \sup_{y' \in B_{3n}} P_{y'}(\tau^{(n)} < \tau_0^{(n)}) \quad \text{ (A.8)}
\]

\[ \leq \sup_{y' \in B_{3n}} P_{y'}(\tau_0^{(n)} > T_1) + \sup_{y' \in B_{3n}} P_{y'}(\tau^{(n)} \leq \ell T_1). \]

For arbitrary \( \hat{y}^{(n)} \in B_{3n}^{(n)} \) converging to some \( \hat{y}, P_{\hat{y}^{(n)}}(\tau_0^{(n)} \leq T_1) \to 1, \) as \( n \to \infty. \) Moreover, with \( \hat{y} \) solving (1.1) for \( \hat{y}(0) = \tilde{y} \) and with \( \hat{t}_1 \) being the length of time it takes \( \hat{y} \) to reach \( \text{cl} \ B_{\eta/2} \),

\[ P_{\hat{y}^{(n)}}(\tau_{\epsilon_0}^{(n)} > T_1) \leq P_{\hat{y}^{(n)}}(\tau_0^{(n)} > \hat{t}_1) \leq P_{\hat{y}^{(n)}}(Y^{(n)}(\hat{t}_1) \notin B_{\eta}) \leq P_{\hat{y}^{(n)}}(|Y^{(n)}(\hat{t}_1) - \hat{y}(\hat{t}_1)| > \frac{\eta}{2}) \]

so that

\[
\limsup_{n \to \infty} \frac{1}{n} \ln P_{\hat{y}^{(n)}}(\tau_{\epsilon_0}^{(n)} > T_1) \leq -\inf(\mathcal{I}_{\hat{y}}'(y') : |y'(\hat{t}_1) - \hat{y}(\hat{t}_1)| \geq \frac{\eta}{2}) < 0. \quad \text{(A.9)}
\]

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

\[ \sup_{y' \in B_{3n}} P_{y'}(\tau_{\epsilon_0}^{(n)} > T_1) \leq e^{-n\chi}. \]

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

\[ \sup_{y' \in B_{3n}} P_{y'}(\tau_{\epsilon_0}^{(n)} > \ell T_1 | \tau_0^{(n)} > (\ell - 1)T_1) \leq \sup_{y' \in B_{3n}} P_{y'}(\tau_0^{(n)} > T_1) \]

so that

\[
\sup_{y' \in B_{3n}} P_{y'}(\tau_{\epsilon_0}^{(n)} > \ell T_1) \leq e^{-n\chi}. \quad \text{(A.10)}
\]

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

\[
\limsup_{n \to \infty} \frac{1}{n} \ln P_{\hat{y}^{(n)}}(\tau^{(n)} \leq \ell T_1)
\]
On the metastability

$$\limsup_{n \to \infty} \frac{1}{n} \ln P_{y(n)}(Y^{(n)}(t) \notin D \text{ for some } t \leq \ell T_1)$$

$$\leq -\inf_{y \in F} I^{*}_{y}(y) = -I^{*}_{\tilde{y}}(\tilde{y}) \cdot \tag{A.11}$$

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

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

Thus, for all $n$ great enough,

$$\sup_{y' \in B^{(n)}_{\eta/2}} P_{y'}(\tau^{(n)} \leq \ell T_1) \leq e^{-n(U-\kappa)}. \tag{A.13}$$

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

$$1_{\{\sigma^{(n)}_i \leq \tau^{(n)}\}} P_{y'_i}(\tau^{(n)} < \tau^{(n)}_i | F^{(n)}(\sigma^{(n)}_i)) \leq 2e^{-n(U-\kappa)}. \tag{A.14}$$

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

$$\limsup_{n \to \infty} \frac{1}{n} \ln P_{y(n)}(\tau^{(n)}_0 > T_1) \leq -\inf(I^{*}_{y}(y') : |y'(\ell) - y(\ell)| \geq \frac{\eta}{2}) < 0,$$

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

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

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, $P_{y(n)}(\tau^{(n)}_0 > T_1) \vee P_{y(n)}(\tau^{(n)} \leq T_1) \leq e^{-nx}$.

Therefore,

$$P_{y(n)}(\tau^{(n)} < \tau^{(n)}_0) \leq P_{y(n)}(\tau^{(n)}_0 > T_1) + P_{y(n)}(\tau^{(n)} \leq T_1) \leq 2e^{-nx}. \tag{A.15}$$

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

$$P_{y(n)}(\tau^{(n)} < e^{n(U-3\kappa)} \leq 4e^{-nx} + P_{y(n)}(e^{-n(U-2\kappa)} \tau^{(n)}_{e^{n(U-2\kappa)}}) < e^{-nx}). \tag{A.16}$$
Since
\[ \tau^{(n)}_{\lfloor n(U - 2\kappa) \rfloor} \geq \sum_{i=1}^{\lfloor n(U - 2\kappa) \rfloor} (\tau^{(n)}_i - \sigma^{(n)}_i) \land 1, \]
by the strong Markov property,
\[ E_{y^{(n)}}((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1|\mathcal{F}^{(n)}(\sigma^{(n)}_i)) = E_{Y^{(n)}}(\sigma^{(n)}_i) (\tau^{(n)}_0 \land 1) \geq \inf_{y' \in B^{(n)}_{3n} \setminus B^{(n)}_{2n}} E_{y'}(\tau^{(n)}_0 \land 1), \]
so, assuming \( n \) is great enough,
\[
P_{y^{(n)}}(e^{-n(U - 2\kappa)} \tau^{(n)}_{\lfloor n(U - 2\kappa) \rfloor} \leq e^{-n\kappa}) \]
\[ \leq P_{y^{(n)}}(e^{-n(U - 2\kappa)} \sum_{i=1}^{\lfloor n(U - 2\kappa) \rfloor} (\tau^{(n)}_i - \sigma^{(n)}_i) \land 1 \leq e^{-n\kappa}) \]
\[ = P_{y^{(n)}}(e^{-n(U - 2\kappa)} \sum_{i=1}^{\lfloor n(U - 2\kappa) \rfloor} ((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1 - E_{y^{(n)}}((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1))) \]
\[ \leq e^{-n\kappa} - e^{-n(U - 2\kappa)} \sum_{i=1}^{\lfloor n(U - 2\kappa) \rfloor} E_{y^{(n)}}((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1)) \]
\[ \leq P_{y^{(n)}}(e^{-n(U - 2\kappa)} \sum_{i=1}^{\lfloor n(U - 2\kappa) \rfloor} ((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1 - E_{y^{(n)}}((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1))) \]
\[ \leq e^{-n\kappa} - \frac{1}{2} \inf_{y' \in B^{(n)}_{3n} \setminus B^{(n)}_{2n}} E_{y'}(\tau^{(n)}_0 \land 1). \]

Let \( T' > 0 \) denote the infimum of time lengths it takes a solution of (1.1) to get from a point in \( B^{(n)}_{3n} \setminus B^{(n)}_{2n} \) to a point in \( B_{y'} \). Since \( Y^{(n)} \) started at point \( y^{(n)} \) started at \( y^{(n)} \in B^{(n)}_{3n} \setminus B^{(n)}_{2n} \) such that \( y^{(n)} \to y' \) converges to the solution of (1.1) started at \( y' \) locally uniformly, \( \tau^{(n)}_0 \) is greater than \( T'/2 \) with great probability, for \( n \) great enough. Therefore, \( \lim_{n \to \infty} E_{y^{(n)}}(\tau^{(n)}_0 \land 1) \geq (T'/2) \land 1 \). Hence, \( e^{-n\kappa} - (1/2) \inf_{y' \in B^{(n)}_{3n} \setminus B^{(n)}_{2n}} E_{y'}(\tau^{(n)}_0 \land 1) < 0 \), for all \( n \) great enough, so that, on recalling that the \( \tau^{(n)}_i - \sigma^{(n)}_i \), for \( i \in \mathbb{N} \), are independent, by Chebyshev’s inequality,
\[
P_{y^{(n)}}(e^{-n(U - 2\kappa)} \sum_{i=1}^{\lfloor n(U - 2\kappa) \rfloor} ((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1 - E_{y^{(n)}}((\tau^{(n)}_i - \sigma^{(n)}_i) \land 1))) \]
\[
\leq e^{-n\kappa} - \frac{1}{2} \inf_{y' \in B_n \setminus B_{2n}} E_y \left( (\tau_0^{(n)} \land 1) \right)
\]
\[
\leq P_y^{(n)} \left( e^{-n(U-2\kappa)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)} \rfloor} \left( (\tau_i^{(n)} - \sigma_i^{(n)}) \land 1 - E_y^{(n)} \left( (\tau_i^{(n)} - \sigma_i^{(n)}) \land 1 \right) \right) \right)
\]
\[
\geq \frac{(T'/2) \land 1}{2} - e^{-n\kappa}
\]
\[
\leq \frac{e^{-2n(U-2\kappa)}}{\left( \left( \frac{(T'/2) \land 1}{2} - e^{-n\kappa} \right)^2 \right)} \sum_{i=1}^{\lfloor e^{n(U-2\kappa)} \rfloor} \text{Var}_{y^{(n)}} \left( (\tau_i^{(n)} - \sigma_i^{(n)}) \land 1 \right)
\]
\[
\leq \frac{e^{-n(U-2\kappa)}}{\left( \left( \frac{(T'/2) \land 1}{2} - e^{-n\kappa} \right)^2 \right)}
\]
Thus, the righthand side of (A.16) converges to 0 so that (A.2) has been proved.

References


